DOCUMENT RESUME

ED 175 906 TM 009 468

AUTHOR Weber, Margaret B.; Argo, Jana K.
TITLE A Study of the Homogeneity of Items Produced From Item Forms Across Different Taxonomic Levels.
PUB DATE Feb 79
NOTE 135p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (Kiawah Island, South Carolina, February, 1979); Best copy available

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Achievement Tests; Cognitive Objectives; *Criterion Referenced Tests; Diabetes; *Dietetics; Graduate Students; Higher Education; *Item Analysis; Knowledge Level; Multiple Choice Tests; Norm Referenced Tests; Nutrition; Taxonomy; *Test Construction; *Test Items
IDENTIFIERS Domain Referenced Tests; *Item Forms; Item Homogeneity; Kidney Disease

ABSTRACT

This study determined whether item forms (rules for constructing items related to a domain or set of tasks) would enable naive item writers to generate multiple-choice items at three taxonomic levels--knowledge, comprehension, and application. Students wrote 120 multiple-choice items from 20 item forms, corresponding to educational objectives related to diabetes mellitus and renal disease. The items were then administered to 20 graduate-level dietetic interns. Results were analyzed by a two-way mixed effects factorial design with one observation per cell. No significant differences between test items were reported for nine of the 20 forms. A comparison of a chi-square index of item homogeneity with the analysis of variance indicated agreement between the procedures in 18 of the 20 item forms. Homogeneous items shared the following characteristics: short item stems, distractors, and correct responses; an average of 10 replacement set structures within item forms; and high average item difficulty indices. Results did not support the use of item forms in generating homogeneous items. (A history of norm-referenced, criterion-referenced, and domain referenced testing; the item forms; and the diabetes and renal disease tests are included.) (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED175906

A STUDY OF THE HOMOGENEITY OF ITEMS PRODUCED FROM ITEM FORMS ACROSS DIFFERENT TAXONOMIC LEVELS

Margaret B. Weber
Assistant Professor
Division of Educational Studies
Emory University

Jana K. Argo
Assistant Professor
Division of Allied Health
Emory University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jana K. Argo

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Eastern Educational Research Association Annual Meeting
Kiawah Island, South Carolina
February, 1979

TM009 468

Table of Contents

	Page
List of Tables and Figures	iii
Abstract	vi
Introduction	1
Review of Related Literature	3
Historical Development of Measurement	3
Norm-Referenced Testing	7
Criterion-Referenced Testing	11
Domain-Referenced Testing	14
Domain-Referenced Test Development	15
Method	24
Procedure	24
Subjects	28
Design	29
Results and Discussion	31
Phase One	31
Phase Two	31
Phase Three	34
Phase Four	34
Phase Five	34
Implications	65
Limitations	67

Page

Conclusion 70

References 73

Appendixes 78

Appendix A 78

Appendix B 81

Appendix C 94

LIST OF TABLES AND FIGURES

		Page
Table 1	Number of Behavioral Objectives by Subject and Taxonomic Level	25
Figure 1	Item Form Packet Arrangement	27
Figure 2	Generalized Design	30
Table 2	Number of Item Forms by Subject Area and Taxonomic Level	32
Table 3	Percent of Agreement between Subject Matter Specialists and Original Item Form Classification	33
Table 4	Test Items by Subject Area and Taxonomic Level	35
Table 5	Percent of Agreement between Subject Matter Specialists and Original Test Item Classification	36
Table 6	Analysis of Variance for Item Form One at the Knowledge Level in Diabetes	38
Table 7	Analysis of Variance for Item Form Two at the Knowledge Level in Diabetes	39
Table 8	Analysis of Variance for Item Form Four at the Knowledge Level in Diabetes	40
Table 9	Analysis of Variance for Item Form Two at the Comprehension Level in Diabetes	41
Table 10	Analysis of Variance for Item Form One at the Knowledge Level in Renal Disease	42

		Page
Table 11	Analysis of Variance for Item Form Three at the Knowledge Level in Renal Disease	43
Table 12	Analysis of Variance for Item Form Four at the Knowledge Level in Renal Disease	44
Table 13	Analysis of Variance for Item Form Three at the Comprehension Level in Renal Disease	45
Table 14	Analysis of Variance for Item Form Two at the Application Level in Renal Disease	46
Table 15	Analysis of Variance for Item Form Three at the Knowledge Level in Diabetes	47
Table 16	Analysis of Variance for Item Form One at the Comprehension Level in Diabetes	48
Table 17	Analysis of Variance for Item Form Three at the Comprehension Level in Diabetes	49
Table 18	Analysis of Variance for Item Form One at the Application Level in Diabetes	50
Table 19	Analysis of Variance for Item Form Five at the Knowledge Level in Renal Disease	51
Table 20	Analysis of Variance for Item Form Two at the Comprehension Level in Renal Disease	52
Table 21	Analysis of Variance for Item Form Five at the Knowledge Level in Diabetes	53
Table 22	Analysis of Variance for Item Form Two at the Application Level in Diabetes	54

Page

Table 23 Analysis of Variance for Item Form Two at the Knowledge Level in Renal Disease 55

Table 24 Analysis of Variance for Item Form One at the Comprehension Level in Renal Disease 56

Table 25 Analysis of Variance for Item Form One at the Application Level in Renal Disease 57

Table 26 Difficulty Indices of Test Items by Subject Matter Area and Taxonomic Level of Item Form 61

Abstract

As an alternative to criterion-referenced tests, which have traditionally limited the inferences that can be drawn from item score to domain score, domain-referenced tests utilizing item forms have been introduced. However, very few attempts have been made to study the characteristics of items generated from item forms and research has failed to support the utility and practicality of item forms.

A study of the homogeneity of items produced from item forms across different taxonomic levels was conducted at a major southern university. Two groups of twenty graduate students in dietetics were recruited for participation in the study.

Twenty behavioral objectives were written from the subject areas of diabetes and renal disease. One item form, suitable for multiple-choice items, was generated for each objective. The 20 item forms were randomly assigned to twenty dietetic graduate students who were unskilled in test construction. A total of 120 test items were generated by the students and consisted of knowledge, comprehension and application items. Items

were divided by subject area into two 60 item tests and were administered to twenty dietetic interns.

Results were analyzed using a two-way mixed effects factorial design with one observation per cell. No significant differences between test items were reported for nine of the twenty item forms. A comparison of a chi-square index of item homogeneity with the analysis of variance procedure indicated agreement between the procedures in 18 of the 20 item forms.

Findings indicated that item forms which produced homogeneous items had the following characteristics: (1) the item stems were short with the majority containing 20 words or less, (2) the length of the distractors and correct responses were shorter than for those found to be non-homogeneous, (3) the number of replacement set structures varied from four to 12, with the majority having 10, (4) the item forms were fairly narrow in breadth, (5) the majority of item forms produced items with high average item difficulty indices.

Findings from this study suggest that further research is required to confirm the use of item forms in the generation of homogeneous test items.

Introduction

Stimulated by the competency-based education movement, the emphasis on behavioral objectives, the sequencing and individualization of instruction, and concerns of cultural bias in norm-referenced measurement, criterion-referenced measurement has experienced a rapid growth during the 1970's. However, considerable inconsistencies exist with respect to definitions of criterion-referenced measurement.

Many sources view criterion-referenced tests as those comprised of items keyed to a set of behavioral objectives. However, recently, criticism has been directed towards these "objectives-referenced" tests as no inference can be drawn from item score to domain score. The item score on an objectives-referenced test only allows interpretation of an examinee's performance on the behaviors specified in the objectives. Items from objectives-referenced tests are not considered to be representative of any behavior domain. Hively (1974) suggested domain-referenced tests as an alternative to objectives-referenced tests which, in contrast, enable one to draw such inferences through the generation of

sets of equivalent related items. In order to maintain item equivalence, Hively suggested the use of item forms. The original item forms developed by Hively (1963) constitute rules for generating a domain of items. Basic notions underlying domain-referenced tests are generalization and transfer. These are essential to the development of legitimate criterion-referenced tests.

The problems of generalization and transfer are essentially those which plague all individuals involved in test construction. In addition, one of the major problems facing state-wide testing programs is the expense of continuous generation of parallel test forms. At the present time, both manual generation of a domain of quality items and computer assistance for the purpose of item generation are difficult, time-consuming and expensive. If unskilled item writers were able to generate equivalent or homogeneous items when structured by item forms, the problems of generalization and transfer as well as the cost of item generation for parallel forms would be reduced. Very few attempts have been made in the past to study the characteristics of items generated from item forms, and research to date has failed to support the utility and practicality of item forms.

Review of Related Literature

Historical Development of Measurement

Measurement, in a broad sense, may be described as the assignment of numerals to objects or events according to rules (Stevens, 1951). People have always been interested in the measurement of human attributes, and embryonic studies can be traced even to ancient China (Nunnally, 1970). However, it was only approximately one hundred years ago that the first systematic attacks were made on problems of measurement.

A fortunate mishap stirred interest in the problem of human observation in 1796 at the observatory at Greenwich, England (Nunnally, 1970). An astronomer's assistant, named Kinneybrook, found himself consistently at odds with his superior concerning observations of the time of transit for certain stars. Astronomers as well as other scientists became interested in Kinneybrook's difficulties. Astronomers began to compare their measurements with one another and found that there were consistent differences between them. Kinneybrook's experiences triggered a recognition that people differ in their judgments and that such individual differences

can be measured and accounted for in scientific work. Following the work of Kinneybrook, Gauss, in 1809, described a thoery of individual differences on observation errors. In 1816 Bessel, an astronomer at the Konigsberg observatory, became interested in measuring the differences between observations reported by trained observers (Nunnally, 1970).

Following Bessel's work, interest developed quickly concerning the psychological explanations of individual differences. Research focused on questions about the threshhold of awareness, i.e., that point at which a visual object, or stimulus of any kind, comes into awareness (Nunnally, 1970). During the latter part of the nineteenth century many scientists from a variety of different specialties studied perceptual acuity.

In a departure from the emphasis on perceptual acuity, in the latter part of the nineteenth century, Sir Francis Galton began to apply the principles of variation and selection in relation to the measurement of physical and mental traits (Mehrens and Lehmann, 1978). A unifying factor in Galton's numerous and varied research activities was his interest in human heredity. In the course of his investigations on heredity, Galton realized the need for measuring the characteristics of

related and unrelated persons and was instrumental in inducing a number of educational institutions to keep systematic anthropometric records on their students. He came to believe that most important personal characteristics are inherited, not only the prominent physical characteristics, but also abilities and personality characteristics. In response to the need for measuring human characteristics in terms of individual differences, Galton invented the "test" (Anastasi, 1976).

Following the work of Galton, the last two decades of the nineteenth century saw a considerable expansion of studies of individual differences in England and the United States. Prominent in this work was James McKeen Cattell, who studied with Galton and then undertook investigations in the United States. The test continued to be largely of a sensory and motor type, measuring sensory speed and reaction time. In an article written by Cattell in 1890, the term "mental test" was used for the first time (Anastasi, 1976). This article described a series of tests that were being administered annually to college students in an effort to determine their intellectual level.

In addition to the work of Cattell, other psychologists began to assemble a number of test series.

Kraepelin (1895) prepared a series of tests to measure what he regarded as basic factors in the characterization of an individual. A German psychologist, Ebbinghaus (1897) administered tests of arithmetic computation, memory span, and sentence completion to school children. In an article published in France in 1895, Binet and Henri criticized most of the available test series as being too largely sensory and as concentrating unduly on simple, specialized abilities (Anastasi, 1976).

Alfred Binet and his co-workers devoted many years to active research on ways of measuring intelligence. In 1904, the Minister of Public Instruction appointed Binet to assist administrative authorities in examining school children suspected of mental deficiency and recommended for transfer from the ordinary school to special classes (Burt, 1921). It was in connection with the objectives of this commission that Binet, in collaboration with Simon, prepared the first Binet-Simon Scale (Binet and Simon, 1905). This scale was composed of 30 items arranged in ascending order of difficulty and was referred to as the Measuring Scale of Intelligence. In the second, or 1908 scale, the number of tests were increased, some unsatisfactory tests from the earlier scale were eliminated, and all tests were grouped into

age levels on the basis of the performance of about 300 normal children between the ages of 3 and 13 years (Anastasi, 1976). The 1908 Measuring Scale of Intelligence was a more refined product than the original scale.

Norm-Referenced Testing

The Binet-Simon test aroused considerable interest in other countries, in England, and especially in the United States. Several translations and revisions of the tests were made in the United States (Nunnally, 1970). Whether intelligence tests are used to screen learners into specialized programs or whether achievement tests are used to indicate learning progress, the impact of the test which discriminates among learners, i.e., the norm-referenced test, is widely acknowledged. The norm-referenced approach allows transformation of the raw score into a scale score. According to this view, the role of the measuring device is to discriminate among learners with respect to the characteristic being measured.

Mehrens and Lehmann (1978) present several uses for norm-referenced measurement. Little argument exists over the use of norm-referenced measurement in aptitude testing where one wishes to make differential predictions.

The second use of norm-referenced measurement deals with the extrinsic rewards of learning in that students who have learned more than others deserve recognition for such accomplishments. In addition, norm-referenced measurement is helpful in fixed-quota selection decisions where there is a limited number of openings. For example, if there is a limited number of openings in a pilot training school, the school would want to select the best of the applicants.

In the United States, many norm-referenced intelligence tests are available including the revised Stanford-Binet, which is currently in widespread use (Mehrens and Lehmann, 1978). The Stanford-Binet, and others patterned after the work of Binet and Simon, are characterized as individual tests as they are used to test only one subject at a time. Many of the tests in these scales require oral responses from the subject. Another characteristic of the Binet-type of test is that it requires a highly trained examiner. For these and other reasons, such tests are not adapted to group administration (Anastasi, 1976).

Group testing, like the first Binet Scale, was developed to meet a pressing practical need. The need for group tests which could be administered by less

highly trained examiners, which could be scored easily, and which could be administered to a large group of people at one time, was realized during World War I. The United States Army was faced with testing over one-half million men in order to classify recruits with respect to intelligence. In response to this need, Army psychologists developed the Army Alpha and Army Beta group intelligence tests.[1]

The evolution of testing from the measurement of intelligence suggested by Galton has led to much interest and concern for individual and group testing. As a result of this interest, and in response to the demand for measurement instruments, test publishing has become a large enterprise. The majority of these tests employ traditional measurement procedures in terms of a norm-referenced approach designed to discriminate among individuals.

Since the mid to late 1960's, traditional achievement testing has been the subject of considerable

---

[1]The Army Alpha test consisted of a series of tasks, such as sequences of instructions, counting, arranging words in sentences, and analogies. The Army Beta test was developed at the same time but was designed for illiterates and those unable to understand English. The Beta test required the subjects to solve geometrical puzzles and to analyze pictures (Nunnally, 1970, 43).

criticism (Geisinger, 1976). The purpose of testing, as it is most frequently used in the existing education system, is primarily the grading and classifying of students in terms of the norm-referenced approach. Critics of the traditional method of achievement testing point out that evaluation procedures are used to make critical and often irreversible decisions about each student's worth and future in the education system. One of the major vehicles used to limit the number of students who have access to advanced education is norm-referenced measurement. Opponents of the norm-referenced approach suggest that the classification of students frequently affects students' entire careers (Bloom, 1964; Hicklin, 1962). Many individuals opposing the classification system feel that continual labeling has an unfavorable effect on many students' self-concept and may have detrimental effects on personality and character development.

In contrast to the notion of having schools classify individuals is the view that education has as its primary function the development of the individual (Bloom, et al., 1971). There is growing concern among educators today that norm-referenced tests do not complement teaching or learning (Glaser and Nitko, 1971).

Criterion-Referenced Testing

Learning may be defined as the acquisition of behaviors brought about by the school environment, instruction and the educational system (Glaser and Nitko, 1971). Therefore, the primary task of educators may be seen as decision-making in regard to how much they want the student to change and how they can assist in that change.

Tyler (1970) describes the function of evaluation as determining the extent to which students change in relation to specified criteria. If the role of education is to produce changes in learners, then someone must decide what changes are possible and desirable. In all of Tyler's work, the importance of clearly defined objectives is evident. Statements of instructional objectives serve as a first step in curricular development, instructional planning and in the evaluation of instruction (Tyler, 1970). Objectives have been a feature of American education since early in this century. If the task of testing is to provide information for making decisions about learning, it is appropriate to compare learning against standards or criteria, which are necessary elements of objectives. With this viewpoint, criterion-referenced tests may provide valuable information.

Glaser (1963) and Popham and Husek (1969) were the first to introduce and popularize the field of criterion-referenced testing. Mehrens and Lehmann (1978) suggest that the support for criterion-referenced measurement has originated in large part from the following: the emphasis on behavioral objectives, the sequencing and individualization of instruction, the development of programmed materials, the increased interest in certification, and the belief that norm-referenced measurement promotes unhealthy competition and is injurious to low-scoring students' self-concepts. In contrast to the use of norm-referenced measurements, criterion-referenced measurement provides information in terms of specific tasks performed by the examinee and produces scores comparing the examinee to some criterion (Denham, 1975).

Along with the recent popularity of criterion-referenced measurement and in response to the increased interest in certification, the era of the 1970's has brought about competency-based education. In September, 1972, the Oregon State Board of Education passed new minimum graduation requirements for students entering ninth grade in the fall of 1974 (Spady, 1977). In order to receive a diploma, students must master minimum standards in three competency areas: personal

development, social responsibility and career development. In addition, students must pass the normal high school courses. Preceding the passage of the new requirements, two years of public hearings explored the alleged inability of the prevailing school curricula, practices and standards to assure graduates could survive socially and economically as independent adults. The 1972 Oregon decision has set in motion a series of actions all across the nation. By 1984, educators believe that most states will have instituted minimal competency standards for their high school graduates (Brodinsky, 1977).

Although competency-based education has gained momentum across the country, there is a lack of agreement among educators as to what constitutes competence and how to describe it (Butler, 1978). Spady (1977) presents a definition of competency-based education based on a concensus of a number of acknowledged experts in the field: a data-based, adaptive, performance-oriented set of integrated processes that facilitate, measure, record and certify within the context of flexible time parameters, the demonstration of known, explicitly stated, and agreed upon learning outcomes that reflect successful functioning in life roles.

Considerable inconsistencies in educational vocabulary exist with respect to criterion-referenced measurement (Alkin, 1974; Donlan, 1974; Millman, 1973, 1974). Popham and Husek (1969) view criterion-referenced measurement as providing the kind of test score information needed to make a variety of individual and programmatic decisions arising in objectives-based instructional programs. Ivens (1972) offers a similar definition: a criterion-referenced test (CRT) is "one comprised of items keyed to a set of behavioral objectives" (2). Harris and Steward (1971) present a much more restrictive definition:

> A pure criterion-referenced test is one consisting of a sample of production tasks drawn from a well-defined population of performances, a sample that may be used to estimate the proportion of performance in that population at which the student can succeed. (2)

Domain-Referenced Testing

The Harris and Stewart definition of a CRT parallels Millman's definition of a domain-referenced test (DRT) as he defines a DRT as "any test consisting of a random or stratified sample of items selected from a well-defined set or class of tasks (a domain)" (Millman, 1974, 315). Similarly, Hively (1974) describes DRTs as contrasted to

a common form of CRT, the objectives-referenced test (ORT), in terms of score generalizability to a domain:

> Inference form item score to domain score is primitive . . . it only tells you about the probability that the student will respond correctly to the same item if you present it again. If you want stronger inferences, you can construct more items for each objective, and then you can sample some of them and estimate the probability that the individual or group will respond correctly to the others . . . That is the only difference between a domain-referenced test and an objectives-referenced test. The strength of the inference depends on the representativeness of the set of items associated with each objective. (104)

Baker (1974) and Osburn (1968) agree with the limitations of objectives-referenced tests identified by Hively. In contrast to single items to measure objectives, domains require the objective-maker to focus on a range of eligible content to which the learner's skill is to apply. The idea behind domain-referenced testing is to have a large universe of content in mind with the implicit objective of generalizing over the entire content domain.

The utility of domain-referenced tests in terms of score generalizability is clear. However, mechanisms for developing DRTs are not widely known nor well accepted.

Domain-Referenced Test Development

In order to operationalize domain-referenced testing,

Hively (1974) defines several terms. Items are defined as "sets of instructions telling how to evoke, detect and score a specific bit of human performance" (6). Items must include directions for presenting the stimuli, recording the response and deciding whether the response is appropriate. A domain is defined as "any clearly defined set of items" (8). Baker (1974) states that domains for testing and teaching represent an attempt to find a reasonable compromise between vagueness and over-precision.

Central to the problem of specifying domains is the definition of content rules or limits (Baker, 1974). Such rules specify the common characteristics of eligible examples, problems, or more generally, stimuli, which will be presented to the learner in the testing and teaching situations. By utilizing rules, domains prohibit the production of trivial objectives and influence the writer in the direction of significant, generalizable skills.

Hively (1974) suggests that a useful way to generate domains is to ask what part of an item can be changed to create other items that test the same general abilities. The permissible replacements for the variable elements can then be listed and a set of substitution rules

enabling the test maker to generate this set of related items can be written. Hively (1974) defines this list of rules for generating a set of related items as an item form.[2] However, the concept of an item form was first developed by Wells Hively II (1963) as an approach to defining criterion behavior for programmed instruction.

The original item forms developed by Hively (1963) consist of an item shell which provides the general framework within which the item content is to be presented, and the replacement set structure, which consists of the rules for generating the content to be used with the item form shell. Combining an item form shell with a replacement set structure gives an item form which constitutes rules for generating a domain of items.

Since the work of Hively, various articles have appeared in the literature with guidelines for the development of item forms. Baker (1974) and Popham (1975) suggest steps for the development of item forms, which they refer to as domain specifications. Each of these procedures is intended to increase the probability of transfer and generalizability.

---

[2]An item form is a process which is capable of generating items with a fixed syntactical structure which contains one or more variable elements and defines a class of item sentences by specifying the replacement sets for the variable elements (Hambleton, et al., 1978, 33).

Baker (1974) presents six components of domain specifications: (1) the development of domain descriptors which could be behavioral objectives, (2) the development of content limits, defined as a set of rules (or a list) of content eligible for inclusion in the test items, (3) the statement of criteria for the constructed responses or the distractors for selected responses, (4) the format, or a description of the form in which the item will be presented to the student, (5) directions which provide the wording that must be comprehendable to the student, and (6) the inclusion of items sampled from within the content limits and intended as representative of the class of responses desired.

Popham (1975) provides a simplified version of Baker's guidelines for the development of domain specifications: (1) a general description is prepared which could be a behavioral objective, (2) a sample test item is prepared which serves to clarify the domain of test items and specify the item format, (3) the behaviors included in the domain are indicated, and (4) the characteristics of response alternatives are specified.

Both Bormuth (1970) and Baker (1974) discuss the advantages of the item form analysis approach as compared to traditional achievement test construction. In

traditional methods of test construction, both the content and behavior outlines are up to the discretion of the test maker, and the selection of content may be haphazard due to the lack of rules to go by. Indeed, Bormuth (1970) states that these "techniques place so few constraints on the item writer that it seems improbable that two test writers could duplicate each other's results" (10).

Bormuth (1970) cites the following advantages for the item form analysis approach to achievement test construction:

(1) the test writer has no options in how the item is phrased,

(2) that an item measures the content claimed can be verified,

(3) the logical relevance of an item to the construction can be specified,

(4) variance attributed to the writer's behavioral disposition should be eliminated, and

(5) the application of a definition to the instruction automatically produces a population of items.

Theoretically, the use of the item form analysis approach "ensures that tests made by the operations will be

independently reproducible given only a knowledge of what operations were used to derive the items and their responses" (35).

Baker (1974) and Heines (1975) suggest that item forms in the development of domain-referenced tests facilitate generalization, transfer, and subject matter structure. A major advantage lies in the ability to produce groups of items that are random or stratified random samples of a specified universe of content (Jackson, 1970). Since the procedure is systematic and rule-bound, it has proven amenable to automation.

Generalization was examined in a study by Macready and Merwin (1973). Consideration was given to the nature of the relationships among items within item forms and how these relationships compare with an ideal case for diagnostic tests in which if a person gets one item within an item form right, one could assume that all items within the item form would be answered correctly. Furthermore, use of any item produced from a given item form would provide information on how a person would respond to all items within that form. This in turn would mean that the item scores on such a test would take on considerable meaning due to their generalizability to other items.

For an item to have the desired relationship with all other items in the set, two conditions would have to be present:

(1) there should be homogeneity between these items

(2) all items within the item form would need to be of equal difficulty.

Macready and Merwin operationally defined a perfectly homogeneous set of items as a group of items which, for a given set of individuals, would yield the probability of unity that any person would pass all items of equal or lesser difficulty than any item which he or she was known to have passed.[3] Thus, an item would be considered inadequate for use in a domain-referenced test if (a) the items within the item form were not homogeneous, (b) the items were not of equal difficulty, or (c) both of the above.

From a study of item forms in multiplication (see Appendix A), Macready and Merwin concluded that in most cases item forms which generate items of moderate difficulty could be used to obtain relatively homogeneous sets of items of equivalent difficulty for a defined

---

[3]This case may be distinguished from a perfectly heterogeneous set of items in which the probability of an individual X, passing any item A is the same whether or not he is known already to have passed any other item, B (Macready and Merwin, 1973, 353).

population of subjects. However, as their study was limited to mathematics, as was the work of Hively (1963, 1974) the utility of item forms in the generation of homogeneous items in other domains is questionable.

The findings of this review indicated the need for more research in the area of domain-referenced testing utilizing the item forms analysis approach. The assumption underlying domain-referenced testing is that scores on such tests describe an examinee's performance on a sample of items and these scores may be generalized to performance on the entire domain of items. Therefore, in order to satisfy the conditions for generalizing, it is essential that items within an item form be comparable as evidenced by empirical methods. In addition, if item forms are written clearly in a precise manner, naive item writers who have had no experience in generating items from item forms, should be able to generate equivalent test items from the same form, given only a knowledge of the operations used.[4] The present study was

---

[4]Equivalent items are those items which show similar item difficulty indices in terms of homogeneity, when examined by empirical procedures. Naive item writers are those persons enrolled as either part-time or full-time students in a Master's Degree program in dietetics who have had no experience in writing items from item forms and who are unskilled in test construction.

conducted to examine the homogeneity of items produced from item forms.

Method

Procedure

The procedures used in the study may be divided into five different phases:

Phase One

During the first phase of the study, the investigator wrote one item form (following Hively, 1974), suitable for multiple-choice test items, for twenty behavioral objectives. These twenty behavioral objectives, also developed by the investigator, were distributed across two subject areas and three taxonomic levels (according to Bloom, 1956) as shown in Table 1.

Phase Two

In order to estimate the content validity of the item forms, three subject matter specialists were asked to match the twenty item forms to the behavioral objectives.

Phase Three

In the third phase of the study, the item forms were sorted into twenty packets of three. Each packet contained two item forms from the knowledge level and one from either the comprehension or application level (as

Table 1

Number of Behavioral Objectives by

Subject and Taxonomic Level

	Subject Area	
Taxonomic Level	Renal Disease	Diabetes
Knowledge	5	5
Comprehension	3	3
Application	2	2

shown in Figure 1). Homogeneity of subject area within the packets was maintained. These twenty packets were randomly assigned to each of twenty graduate students in dietetics who are naive item writers, i.e. those unfamiliar with test construction. Each student was asked to generate two test item stems and correct responses for each of the three item forms. Thus, a total of 120 test items were generated (60 in each subject area). The items generated by this process included 40 knowledge, 12 comprehension and eight application items in each subject area. Three distracters were added to each test item by the investigator and were reviewed by a test construction specialist.[5]

Phase Four

The 120 multiple choice test items were examined for content validity during the fourth stage of the study. The same subject matter specialists who participated in the validity study of item forms were asked to classify the behavior of the test items according to taxonomic level and subject matter. Content validity of the test items was examined in terms of the percent of agreement

---

[5]These individuals generating test items did not generate distracters due to the constraints included in teaching this aspect of test construction.

Figure 1

Item Form Packet Arrangement

K_1	K_3	K_5	K_2	K_4
K_2	K_4	K_1	K_3	K_5
C_1	C_2	C_3	A_1	A_2
K_1	K_3	K_5	K_2	K_4
K_2	K_4	K_1	K_3	K_5
C_1	C_2	C_3	A_1	A_2

between content specialists and students generating the items.

Phase Five

Separation of test items was based on subject matter. Thus, the 120 test items were divided into two 60 item tests. Test items were then randomly ordered within each of the two tests. These tests were administered, in the order that the material was taught, to the twenty dietetic interns during regular class sessions. Each student was allowed sufficient time to complete the test.

Subjects

The subjects who were recruited to participate in this study represent a sample of convenience. Two groups of twenty graduate subjects were utilized during the process of the study. Twenty female students, enrolled in a graduate dietetic program, generated items from the item forms written by the investigator. All of the subjects were less than 35 years of age.

Test items generated by the graduate students were administered to twenty dietetic interns (19 females and 1 male) enrolled in a Dietetic Internship Program. All dietetic interns were less than 30 years of age and all had received undergraduate degrees in nutrition or dietetics.

Design

The design selected as appropriate for this study was simple factorial design with one observation per cell (see Figure 2). Data from each subject area, from each taxonomic level, and from each item form were analyzed separately. However, the hypotheses remained the same for each of the twenty analyses.

Hypothesis

H_1: There is no significant difference between test items.

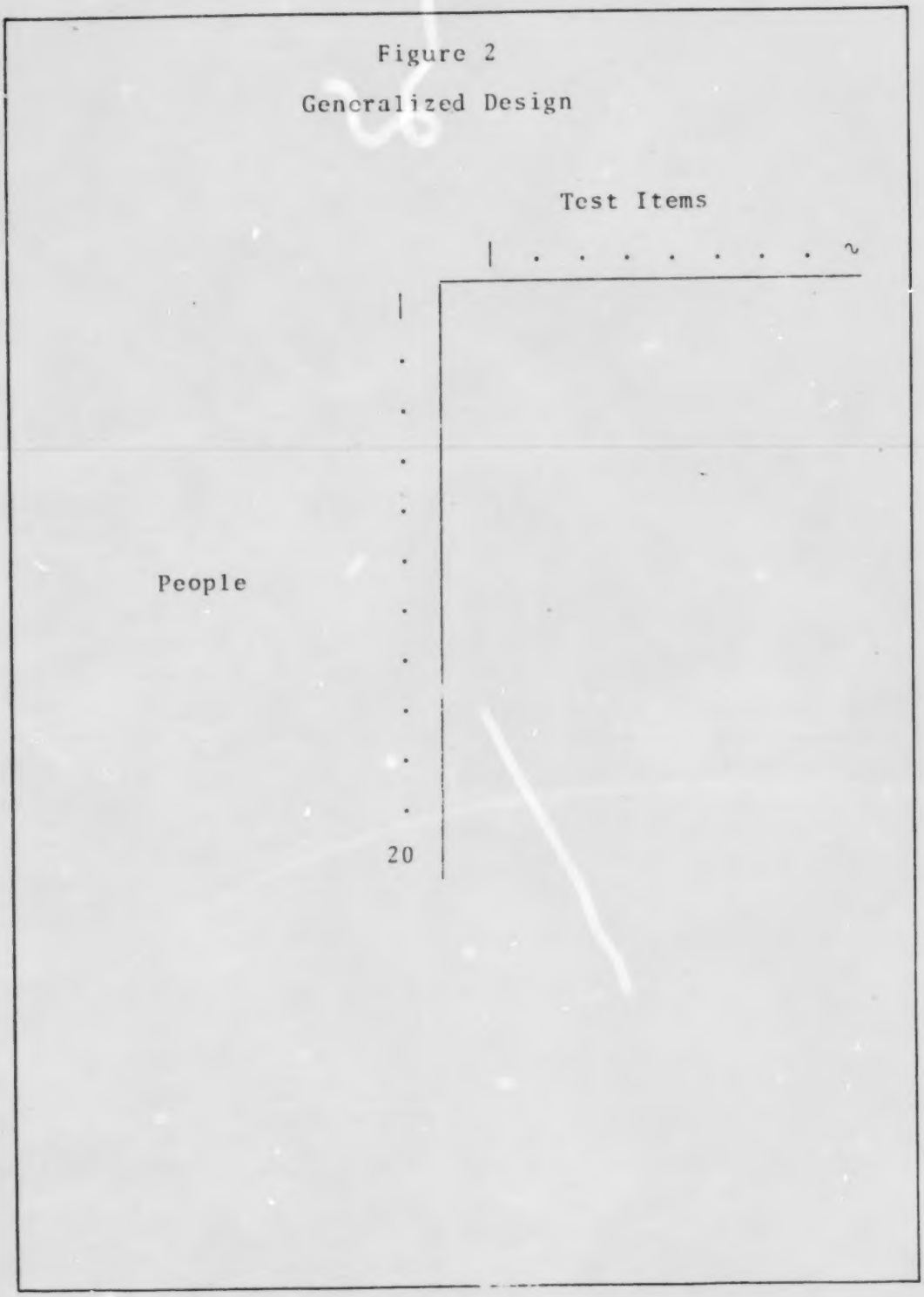
These 20 hypotheses were tested using a two-way mixed effects factorial analysis of variance procedure.[6] In addition, the homogeneity of items within item forms was examined using a revision of Popham's chi-sqaure technique (following Pettie and Oosterhof, 1976).

---

[6]This design is a two-way factorial design with one observation per cell. As such, no estimate of within cells variance estimate is available and no test of interaction can be computed. Thus, this analysis assumes no interaction. The items are a random effect as they are sampled within the parameters defined by each item form. The subjects represent a fixed effect.

Figure 2

Generalized Design



Results and Discussion

Phase One

Twenty behavioral objectives were written from two subject areas in nutrition education. These included knowledge, comprehension and application objectives written in the subject areas of diabetes mellitus and renal disease. Utilizing the twenty behavioral objectives, one item form appropriate for multiple-choice test items was generated for each objective as shown in Table 2. The number of possible test items which could be generated from an item form varied from four to thirty-two, depending on the specificitivity of the item form. In no case were all items possible included within each item form (see Appendix B).

Phase Two

Content validity estimates were obtained from subject matter specialists. Three specialists matched the twenty item forms to the behavioral objectives. Results of the validity estimates indicated agreement between content specialists and the original item form classifications (see Table 3).

Table 2

Number of Item Forms by Subject Area and Taxonomic Level

	Subject Area	
Taxonomic Level	Renal Disease	Diabetes
Knowledge	5	5
Comprehension	3	3
Application	2	2

Table 3

Percent of Agreement between Subject Matter Specialists and Original Item Form Classification

(N = 3)

Subject Matter Specialist	Percent of Agreement	
	Diabetes	Renal
1	100%	100%
2	90%	100%
3	100%	100%

Phase Three

Item forms were duplicated in twenty packets of three forms each. Each packet was randomly assigned to one of twenty naive item writers, i.e., persons unskilled in test construction. Each item writer generated two multiple-choice test items per item form. A total of 120 test items was generated (see Table 4). Appropriate distracters were generated and reviewed by a specialist in test construction.

Phase Four

Content validity estimates for the test items were obtained from the same subject matter specialists who participated in the validity study of item forms (see Table 5).[7]

Phase Five

The 120 test items were divided into two 60 item tests based on subject matter. Test items were randomly

[7]The descrepancy between validity estimates of the item forms and those of the test items should be noted. One possible explanation for the low validity estimates of the test items could be the subject matter specialists' unfamiliarity with the taxonomy. Although all three individuals have practiced in the field of dietetics, none have had experience in test construction.

Table 4

Test Items by Subject Area and Taxonomic Level

	Subject Area	
Taxonomic Level	Diabetes	Renal
Knowledge	40	40
Comprehension	12	12
Application	8	8

Table 5

Average Percent of Agreement between Subject Matter Specialists and Original Test Item Classifications

	Percent of Agreement	
Subject Matter Specialists	Diabetes	Renal
1	73%	47%
2	67%	63%
3	80%	87%

ordered within tests. The two tests were administered to twenty dietetic interns.

Data from each subject area, from each taxonomic level and from each item form were analyzed separately using a two-way factorial design with one observation per cell. No significant differences between items were reported for nine of the twenty item forms. Thus, nine item forms contained homogeneous items.

No significant differences between items are reported for four item forms in the area of diabetes. Of these four item forms, three are at the knowledge level and one at the comprehension level as shown in Tables 6, 7, 8, and 9. The remaining five item forms for which no significant differences between test items are reported are in the area of renal disease. Three of these item forms are at the knowledge level, one is a comprehension level item form, and one is an application level item form (see Tables 10-14).

In five of the nine item forms reporting no significant differences between items, significant nonadditivity is also reported. However, as the items represent a random effect and the subjects a fixed effect, these findings do not influence interpretation of the F ratios. The remaining item forms for which significant differences between items were found are presented in Tables 15-25.

Table 6

Analysis of Variance for Item Form One at the Knowledge Level in Diabetes (8 items)

Source of Variation	df	SS	MS	F	P
Between People	19	5.09	.27		
Within People	100	9.50	.10		
Between Measures	5	.44	.09	.94	.47
Residual	95	9.06	.10		
Nonadditivity	1	1.96	1.96	25.94	.000
Balance	94	7.10	.08		
Total	119	14.59	.12		

$\bar{X}$ = .64 $\bar{X}_{..}$ = .86 Median Item Difficulty = 85.50

S = 1.27 α = .64 χ^2 = 7.51

standardized α = .55 sig. = .500

Table 7

Analysis of Variance for Item Form Two at the Knowledge Level in Diabetes (8 items)

Source of Variation	df	SS	MS	F	P
Between People	19	1.73	.09		
Within People	40	2.00	.05		
Between Measures	2	.03	.02	.32	.73
Residual	38	1.97	.05		
Nonadditivity	1	.22	.22	4.58	.04
Balance	37	1.75	.05		
Total	59	3.73	.06		

$\bar{X}$ = 2.80 $\bar{X}_{..}$ = .93 Median Item Difficulty = 99.70

S = .52 α = .43 χ^2 = 1.39

standardized α = .41 sig.=.975

Table 8

Analysis of Variance for Item Form Four at the Knowledge Level in Diabetes

(8 items)

Source of Variation	df	SS	MS	F	P
Between People	19	6.20	.33		
Within People	100	7.67	.08		
Between Measures	5	.57	.11	1.52	.19
Residual	95	7.10	.07		
Nonadditivity	1	1.21	1.21	19.34	.00003
Balance	94	5.89	.06		
Total	119	13.87	.12		

$\bar{X}$ = 5.20 $\bar{X}..$ = .87 Median Item Difficulty = 90.17

S = 1.40 α = .77 χ^2 = 6.10

standardized α = .74 sig. = .500

Table 9

Analysis of Variance for Item Form Two at the Comprehension Level in Diabetes (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	1.65	.09		
Within People	40	6.00	.15		
Between Measures	2	.40	.20	1.36	.27
Residual	38	5.60	.15		
Nonadditivity	1	.49	.49	3.54	.07
Balance	37	5.11	.14		
Total	59	7.65	.13		

$\bar{X}$ = 2.55 $\bar{X}_{..}$ = .85 Median Item Difficulty = 90.00

S = .51 α = -.70 χ^2 = 4.17

standardized α = -.69 sig. =.250

Table 10

Analysis of Variance for Item Form One at the Knowledge Level in Renal Disease (8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	11.22	.59		
Within People	140	28.75	.21		
Between Measures	7	2.18	.31	1.56	.15
Residual	133	26.58	.20		
Nonadditivity	1	.49	.49	2.49	.12
Balance	132	26.08	.20		
Total	159	39.97	.25		

$\bar{X}$ = 3.90 $\bar{X}_{..}$ = .49 Median Item Difficulty = 50.50

S = 2.17 α = .66 χ^2 = 21.23

standardized α = .66 sig.= .005

Table 11

Analysis of Variance for Item Form Three at the Knowledge Level in Renal Disease (8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	4.99	.26		
Within People	120	18.00	.15		
Between Measures	6	1.24	.21	1.41	.22
Residual	114	16.76	.15		
Nonadditivity	1	.72	.72	5.05	.03
Balance	113	16.04	.14		
Total	139	22.99	.17		

$\overline{X}$ = 5.55 $\overline{X}_{..}$ = .79 Median Item Difficutly = 80.17

S = 1.36 α = .44 χ^2 = 12.72

standardized α = .45 sig. = .050

Table 12

Analysis of Variance for Item Form Four at the Knowledge Level in Renal Disease (8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	8.20	.43		
Within People	100	11.00	.11		
Between Measures	5	.20	.04	.35	.88
Residual	95	10.80	.11		
Nonadditivity	1	.14	.14	1.21	.27
Balance	94	10.66	.11		
Total	119	19.20	.16		

$\bar{X}$ = 4.80 $\bar{X}_{..}$ = .80 Median Item Difficulty = 80.50

S = 1.61 α = .74 χ^2 = 11.20

standardized α = .74 sig.= .100

Table 13

Analysis of Variance for Item Form Three at the Comprehension Level in Renal Disease (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	4.80	.25		
Within People	60	8.00	.13		
Between Measures	3	.90	.30	2.41	.08
Residual	57	7.10	.12		
Nonadditivity	1	1.00	1.00	9.19	.004
Balance	56	6.10	.11		
Total	79	12.80	.16		

$\bar{X}$ = 3.20 $\bar{X}$.. = .80 Median Item Difficulty = 80.00

S = 1.01 α = .51 χ^2 = 5.62

standardized α = .45 sig. = .100

Table 14

Analysis of Variance for Item Form Two at the Application Level in Renal Disease (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	6.75	.36		
Within People	60	12.00	.20		
Between Measures	3	.45	.15	.75	.53
Residual	57	11.55	.20		
Nonadditivity	1	.45	.45	2.29	.14
Balance	56	11.10	.20		
Total	79	18.75	.24		

$\bar{X}$ = 1.50 $\bar{X}$..= .38 Median Item Difficulty = 40.00

S = 1.19 α = .43 χ^2 = 6.25

standardized α = .41 sig.=.500

Table 15

Analysis of Variance for Item Form Three at the Knowledge Level in Diabetes

(8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	3.15	.17		
Within People	80	15.60	.20		
Between Measures	4	4.10	1.02	6.77*	.0001
Residual	76	11.50	.15		
Nonadditivity	1	.67	.67	4.67	.03
Balance	75	10.83	.14		
Total	99	18.75	.19		

*$p \leq .05$

$\bar{X} = 3.75$ $\bar{X}.. = .75$ Median Item Difficulty = 90.50

$S = .91$ $\alpha = .09$ $\chi^2 = 38.92$

standardized $\alpha = .16$ sig. = .005

Table 16

Analysis of Variance for Item Form One at the Comprehension Level in Diabetes (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	2.40	0.13		
Within People	40	9.33	.23		
Between Measures	2	3.43	1.72	11.06*	.0002
Residual	38	5.90	0.16		
Nonadditivity	1	.84	.84	6.16	.02
Balance	37	5.06	0.14		
Total	59	11.73	0.20		

*$p \leq .05$

$\bar{X} = 2.20$ $\bar{X}.. = .73$ Median Item Difficulty = 90.00

$S = .62$ $\alpha = -.23$ $\chi^2 = 29.14$

standardized $\alpha = -.15$ sig.= .005

Table 17

Analysis of Variance for Item Form Three at the Comprehension Level in Diabetes (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	3.14	.17		
Within People	60	7.75	.13		
Between Measures	3	1.04	.35	2.94*	.04
Residual	57	6.71	.12		
Nonadditivity	1	.68	.68	6.27	.02
Balance	56	6.04	.11		
Total	79	10.89	.14		

*$p \leq .05$

$\bar{X} = 3.35$ $\bar{X}_{..} = .84$ Median Item Difficulty = 87.50

$S = .81$ $\alpha = .29$ $\chi^2 = 6.57$

standardized $\alpha = .34$ sig. = .050

Table 18

Analysis of Variance for Item Form One at the Application Level in Diabetes

(4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	4.67	.25		
Within People	40	8.67	.22		
Between Measures	2	2.43	1.22	7.42*	.002
Residual	38	6.23	.16		
Nonadditivity	1	2.31	2.31	21.78	.00004
Balance	37	3.92	.11		
Total	59	13.33	.23		

*$p \leq .05$

$\overline{X} = 2.00$ $\overline{X}.. = .67$ Median Item Difficulty = 75.00

$S = .86$ $\alpha = .33$ $\chi^2 = 27.33$

standardized $\alpha = .06$ sig. = .005

Table 19

Analysis of Variance for Item Form Five at the Knowledge Level in Renal Disease (8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	4.62	.24		
Within People	140	24.88	.18		
Between Measures	7	2.74	.39	2.36*	.03
Residual	133	22.13	.17		
Nonadditivity	1	.83	.83	5.17	.02
Balance	132	21.30	.16		
Total	159	29.49	.19		

* $p \leq .05$

$\bar{X} = 6.05$ $\bar{X}.. = .76$ Median Item Difficulty = 80.50

$S = 1.39$ $\alpha = .32$ $\chi^2 = 19.40$

standardized $\alpha = .31$ sig. = .005

Table 20

Analysis of Variance for Item Form Two at the Comprehension Level in Renal Disease (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	4.40	.23		
Within People	40	7.33	.18		
Between Measures	2	1.63	.82	5.44*	.008
Residual	38	5.70	.15		
Nonadditivity	1	1.25	1.25	10.37	.003
Balance	37	4.45	.12		
Total	59	11.73	.20		

* $p \leq .05$

$\bar{X} = 2.20$ $\bar{X}.. = .73$ Median Item Difficulty = 82.50

$S = .83$ $\alpha = .35$ $\chi^2 = 16.67$

standardized $\alpha = .31$ sig. = .005

Table 21

Analysis of Variance for Item Form Five at the Knowledge Level in Diabetes

(8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	3.72	.20		
Within People	140	32.38	.23		
Between Measures	7	4.44	.63	3.02*	.006
Residual	133	27.93	.21		
Nonadditivity	1	.24	.24	1.15	.29
Balance	132	27.69	.21		
Total	159	36.09	.23		

*$p \leq .05$

$\bar{X} = 2.75$ $\bar{X}.. = .34$ Median Item Difficulty = 25.50

$S = 1.25$ $\alpha = -.07$ $\chi^2 = 110.67$

standardized $\alpha = -.07$ sig.= .005

Table 22

Analysis of Variance for Item Form Two at the Application Level in Diabetes (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	6.30	.33		
Within People	60	13.50	.22		
Between Measures	3	3.70	1.23	7.17*	.0004
Residual	57	9.80	.17		
Nonadditivity	1	.25	.25	1.50	.23
Balance	56	9.55	.17		
Total	7[illegible]	19.80	.25		

*$p \leq .05$

$\bar{X} = 1.80$ $\bar{X}_{..} = .45$ Median Item Difficulty = 52.50

$S = 1.15$ $\alpha = .48$ $\chi^2 = 39.52$

standardized $\alpha = .53$ sig. = .005

Table 23

Analysis of Variance for Item Form Two at the Knowledge Level in Renal Disease (8 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	6.15	.32		
Within People	140	30.25	.22		
Between Measures	7	7.70	1.10	6.49*	.0000
Residual	133	22.55	.17		
Nonadditivity	1	.20	.20	1.17	.28
Balance	132	22.35	.17		
Total	159	36.46	.23		

* $p \leq .05$

$\bar{X} = 5.20$ $\bar{X}.. = .65$ Median Item Difficulty = 65.00

$S = 1.61$ $\alpha = .48$ $\chi^2 = 59.23$

standardized $\alpha = .50$ sig. = .005

Table 24

Analysis of Variance for Item Form One at the Comprehension Level in Renal Disease (4 Items)

Source of Variation	df	SS	MS	F	p
Between People	19	4.05	.21		
Within People	60	13.50	.22		
Between Measures	3	2.55	.85	4.42*	.007
Residual	57	10.95	.19		
Nonadditivity	1	.62	.62	3.38	.07
Balance	56	10.33	.18		
Total	79	17.55	.22		

* $p \leq .05$

$\bar{X} = 2.70$ $\bar{\bar{X}}.. = .68$ Median Item Difficulty = 82.50

$S = .92$ $\alpha = .10$ $\chi^2 = 18.57$

standardized $\alpha = -.03$ sig = .005

Table 25

Analysis of Variance for Item Form One at the Application Level in Renal Disease (4 Items)

Source of Variation	df	SS	MS	F	P
Between People	19	7.30	.38		
Within People	60	12.50	.21		
Between Measures	3	2.70	.90	5.23*	.003
Residual	57	9.80	.17		
Nonadditivity	1	.01	.01	.08	.77
Balance	56	9.79	.17		
Total	79	19.80	.25		

* $p \leq .05$

$\bar{X} = 2.20$ $\bar{X}.. = .55$ Median Item Difficulty = 52.50

$S = 1.24$ $\alpha = .55$ $\chi^2 = 26.19$

standardized $\alpha = .57$ sig. = .005

Characteristics of those item forms which resulted in no significant differences between test items were examined. Such factors as the order of placement of items on the test, the length of the item stem, the length of the correct response and distracters, and the complexity and breadth of the replacement set structures were considered in the examination.

Examination of item forms reporting homogeneous test items reveals that test items are randomly distributed throughout the tests. Thus, the order of placement does not appear to influence homogeneity. The majority of forms producing homogeneous items contain short item stems, i.e. approximately twenty words or less. In comparison, some of the forms producing items which were not homogeneous have item stems of 25 words or more. In relationship to the length of the correct response and distracters, those item forms reported as homogeneous appear to have shorter responses than those shown to lack homogeneity. The number of replacement set structures within item forms with homogeneous items varied from four to twelve, with the majority having approximately ten. In contrast to this finding, it appears that those forms reported as non-homogeneous either have very few (four-eight) or very many (18-32) replacement

set structures. However, in terms of the characteristics examined, the complexity and breadth of the item form in terms of the number of choices necessary for item generation, seem to be most related to the homogeneity of the items. Those items reported as homogeneous contain less replacement set structures resulting in less choices to be made during item generation. In addition, in those cases where the item form produced homogeneous items the forms appear more narrow in breadth. This observation is relatively similar to that obtained by Macready and Merwin (1973).

In addition, it should be noted that six of the nine item forms reporting homogeneous items are at the knowledge level of the taxonomy. This finding may be related to the breadth of the replacement set structure as well as to the degree of structure of the subject matter contained in the homogeneous forms. Hively et al. (1968) showed that item forms analysis procedures were more easily employed with highly structured subject matter areas such as mathematics. In comparison to behavior required at the comprehension and application levels, behavior at the knowledge level appears to be somewhat more highly structured and straight forward in terms of responses required.

Results of the present study also suggest that in most cases item forms which generate items of easy difficulty can be used to obtain relatively homogeneous sets of items. Difficulty indices are presented in Table 26. This observation lends support to the previous discussion concerning the homogeneity of knowledge level items since items written at the knowledge level should produce items with easy difficulty indices as compared to the higher levels of the taxonomy. However, comparisons with findings in the literature reveal that items of moderate difficulty were found to obtain relatively homogeneous sets of items (Macready and Merwin, 1973).

Another and more recent test of item homogeneity was developed by Pettie and Oosterhof (1976). A comparison of Pettie and Oosterhof's chi-square index of item homogeneity with Hoyt's analysis of variance procedure indicates agreement between the procedures in eighteen of the twenty item forms.

The disagreements found in two of the item forms may reflect error introduced in the use of the computed median item difficulty as an estimated parameter. Indeed, this goodness of fit test is less sensitive than is Hoyt's procedure. Therefore, in the two instances of disagreement, it is assumed that the analysis of variance

Table 26

Difficulty Indices of Test Items By Subject Matter Area and Taxonomic Level of Item Form

Taxonomic Level of Item Form	Number of Items on Test	Diabetes Item Difficulties	Number of Items on Test	Renal Item Difficulties
Knowledge #1	14	.95	2	.55
	16	.85	13	.45
	18	.85	16	.55
	25	.90	25	.50
	32	.75	28	.50
	41	1.00	33	.70
	48	.85	37	.35
	60	1.00	58	.70
Knowledge #2	8	.95	4	.65
	12	.95	11	.55
	17	1.00	14	.75
	33	1.00	30	.55
	37	1.00	31	.95
	46	1.00	48	.90
	56	.90	53	.65
	59	1.00	59	.20

Table 26 (continued)

Taxonomic Level of Item Form	Number of Items on Test	Diabetes Item Difficulties	Number of Items on Test	Renal Item Difficulties
Knowledge #3	7	.85	3	.90
	19	.65	20	.80
	24	1.00	24	.80
	26	.40	27	1.00
	28	1.00	34	.60
	43	.85	46	.80
	45	1.00	54	.90
	52	.90	56	.75
Knowledge #4	3	.90	21	.75
	21	.90	23	.80
	27	.90	35	.85
	29	.95	36	1.00
	36	1.00	38	1.00
	44	.80	43	.85
	50	.75	51	.80
	54	1.00	57	.75

Table 26 (continued)

Taxonomic Level of Item Form	Number of Items on Test	Diabetes Item Difficulties	Number of Items on Test	Renal Item Difficulties
Knowledge #5	1	.40	8	.60
	4	.50	10	.80
	6	.15	12	.55
	13	.50	26	.80
	31	.60	39	.95
	35	.25	42	.85
	39	.15	50	.85
	47	.25	55	.65
Comprehension #1	9	.85	19	.70
	15	1.00	40	.40
	23	.40	44	.90
	34	.95	49	.70
Comprehension #2	20	1.00	5	.70
	42	.95	17	1.00
	51	.75	32	.55
	58	.85	47	.95

Table 26 (continued)

Taxonomic Level of Item Form	Number of Items on Test	Diabetes Item Difficulties	Number of Items on Test	Renal Item Difficulties
Comprehension #3	22	.85	7	.80
	49	.90	22	.65
	53	.65	29	.95
	57	.95	60	.80
Application #1	11	.95	1	.65
	30	.55	9	.55
	40	1.00	41	.80
	55	.50	45	.40
Application #2	2	.45	6	.45
	5	.65	15	.40
	10	.10	18	.40
	38	.60	52	.25

describes the presence or absence of homogeneous items with greater power than does the chi-square technique.

In summary, the results of this study suggest the following:

- Item forms written for subject matter content in a highly technical science-related area do not appear to work.
- The order of placement of items on a test does not seem to influence the homogeneity of test items.
- The length of the item stem appears to be related to the homogeneity of test items.
- Homogeneity of items may be related to the length of the correct responses and distracters.
- The complexity and breadth of the item forms appears to be related to the homogeneity of items.
- Homogeneity of items appears to be related to the taxonomic level of the item form.
- The degree of difficulty of the items may be related to homogeneity.

Implications

The results of this study are important for both educators and testing programs involved in the generation of tests. Of particular interest is the possibility of having unskilled item writers generate homogeneous items

which have been structured by item forms. Questions about the appropriateness of item forms in general, across different subject areas and taxonomic levels, should be addressed. Perhaps item forms appropriate for generating items at the higher levels of the taxonomy could result in homogeneous test items. If this is the case, educators should be encouraged to incorporate items which test higher levels of behavior in the cognitive domain.

It appears from this study that the majority of item forms written at higher levels of the taxonomy than the knowledge level, result in more heterogeneous test items. Further research is necessary to determine if this is the case with subject matter areas which are less scientifically oriented.

Further research is also needed to examine the effects of the breadth of the replacement set structures, the length of the item stems, and the length of the correct responses and distractors on the homogeneity of items. In addition, other characteristics of the item form may effect the homogeneity of items in ways which have not been studied.

If an educational goal is to reduce the cost of item generation for parallel test forms and other forms of tests, one vehicle for this goal may be through the use

of item forms. At the present time manual generation of a domain of quality items necessary for parallel test forms is difficult, time-consuming, and expensive. In addition, the utility of computer assistance for test construction has yet to be realized. Consequently, the use of item forms in the production of tests may be the desirable strategy when presently, evidence for the utility of other methods does not exist.

Limitations

The size and composition of the sample, as well as the subject matter, limit the conclusions that can be drawn from this study. In addition, the circumstances of the testing situation, the type of analysis used, and the validity of the testing instrument require refinement.

The number of observations were limited, thereby restricting the sample of student responses to the test items. In addition, more items generated from item forms written at each of the levels of the taxonomy would provide more information concerning the utility of item forms. The subject matter could also have been a limiting factor. Previous studies have suggested that the procedures involved in the generation of items from item forms are more easily employed with highly structured subject areas. Dietetics is a science which is

considered to be a highly specified field. Therefore, studies of the homogeneity of items from item form written in less specified areas may have resulted in different findings. The topics of diabetes and renal disease are also very narrow topics, indicating the possibility that performance in these areas may not generalize to other areas of dietetics or to other educational fields.

Another factor considered as a limitation is the circumstances of the testing situation. Tests were administered as part of the instructional program where students do not receive grades, as opposed to the graduate courses in which students were also enrolled. In addition, tests were administered at the end of a six week period, during which time students were in the classroom setting for approximately six hours each day. Thus, both the fact that students may not have studied for the tests and the degree of fatigue of the students may have influenced responses to test items.

The extent to which the data from the test items were analyzed should also be considered as a possible limitation. The difficulty of the test items, as specifically influenced by the correct responses and distracters, was not considered in the analyses. However,

an effort was made to develop homogeneous responses and distracters although the contribution of these to the variance of items within item forms was not a component of the analyses.

Lastly, the validity of the testing instrument may have been a limitation due to the low content validity estimates obtained. Results of the validity studies indicate the need for individuals familiar with both subject matter content and taxonomic classifications to participate in the validity study.

Conclusion

One of the major problems facing state-wide testing programs is the expense of continuous generation of parallel test forms. In addition, many individuals assuming the roles of educators do not have adequate training in the area of test construction. Therefore, it is important to pursue the possibility of whether unskilled item writers would be able to write equivalent or homogeneous items when structured by item forms.

The purpose of this study was to examine the homogeneity of items produced from item forms across different taxonomic levels within the cognitive domain in an attempt to produce evidence of the utility of item forms in the production of tests. Specifically, the following question was examined:

1. Do item forms enable naive item writers to generate homogeneous items at different taxonomic levels in two areas of nutrition education?

Twenty behavioral objectives were written from the areas of diabetes and renal disease in dietetic education. These included knowledge, comprehension and application

objectives. One item form was generated for each objective. These item forms were then duplicated in twenty packets of three forms each. Each packet was randomly assigned to one of twenty naive item writers. Each item writer generated two test items per item form. A total of 120 test items were generated. Appropriate distracters were generated and items were divided by subject area into two 60 item tests. Content validity estimates were obtained from subject matter specialists. The two tests were administered to twenty dietetic interns.

A two-way mixed effects factorial design with one observation per cell was used to test H_1 of no significant differences between test items. Data from each subject area, from each taxonomic level and from each item form were analyzed separately.

No significant differences between items were reported for nine of the twenty item forms. Of the forms found to be homogeneous, four were in the area of diabetes with the remainder being in the area of renal disease. Three of the item forms resulting in homogeneous items were written at the knowledge level of the taxonomy for both diabetes and renal disease. In addition, one item form at the comprehension level for both diabetes and renal disease resulted in homogeneous items. The one

additional finding of no difference between items occurred at the application level in the area of renal disease. Results of the analysis of variance procedure as compared with those of a chi-square index of item homogeneity indicated agreement between these procedures in 18 of the 20 item forms. Thus, as indicated by the results of this study, further research is necessary to support the use of item forms in the generation of homogeneous test items.

References

Alkin, M. C. "Criterion-referenced measurement" and other such terms. In C. W. Harris, M. C. Alkin, and W. J. Popham (Eds.), Problems in criterion-referenced measurement. CSEmonograph series in evaluation, No. 3. Los Angeles: Center for the Study of Evaluation, Univeristy of California, 1974.

Anastasi, A. Psychological Testing. New York: Macmillan Publishing Company, Inc. 1976.

Baker, E. L. Beyond objectives: Domain-referenced tests for evaluation and instructional improvement. Educational Technology, 1974, 14, 10-16.

Binet, A. and Simon, T. Methodes and nouvelles pour le diagnostic du niveau intellectural des anormaux. Annee Psychologique, 1905, 11, 191-244.

Bloom, B. S. Stability and change in human characteristics. New York: Wiley, 1964.

Bloom, B. S., Hastings, J. T. and Madaus, G. F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill Book Company, 1971.

Bormuth, J. R. On the theory of achievement test items. Chicago: University of Chicago Press, 1970.

Brodinsky, B. Back to the basics: The movement and its meaning. Phi Delta Kappan, 1977, 58, 522-527.

Burt, C. Mental and Scholastic Tests. London: P. S. King and Son, Ltd., 1927.

Butler, R. C. The concept of competence: An operational definition. Educational Technology, 1978, 18, 7-18.

Denham, C. H. Criterion-referenced, domain-referenced and norm-referenced measurement: A parallax view. Educational Technology, 1975, 15, 9-13.

Donlan, T. F. Some needs for clearer terminology in criterion-referenced measurement. Presented at the Annual meeting of the National Council for Measurement in Education, 1974.

Geisinger, K. F. A systems approach to item production and review in a computer managed instruction project. San Francisco, Calif.: American Educational Research Association, 1976. (ERIC Document Reproduction Service No. ED 121 280)

Glaser, R. Instructional technology and the measurement of learning outcomes: some questions. American Psychologist, 1963, 18, 519-521.

Glaser, R. and Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Hambleton, R. K., Swaminathan, H., Algina, J. and Coulson, D. B. Criterion-referenced testing and measurement: A review of technical issues and developments. Review of Educational Research, 1978, 48, 1-47.

Hambleton, R. K. and Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-170.

Heines, J. An examination of the literature on criterion-referenced and computer-assisted testing, 1975. (ERIC Document Reproduction Service No. ED 116 633)

Hicklin, W. J. A study of long-range techniques for predicting patterns of scholastic behavior. Unpublished doctoral dissertation, University of Chicago, 1962.

Hively, W. Defining criterion-behavior for programmed instruction in elementary mathematics. Harvard University, 1963.

Hively, W. Introduction to domain-referenced testing. Educational Technology, 1974, 14, 5-10.

Hively, W., Patterson, H. L. and Page, S. A. A "universe-defined" system of arithmetic achievement tests. Journal of Educational Measurement, 1968, 5, 275-290.

Ivens, S. H. A pragmatic approach to criterion-referenced measures. Chicago, Ill.: National Council on Measurement in Education and AERA, 1972. (ERIC Document Reproduction Service No. ED 064 406)

Jackson, R. Developing criterion-referenced tests. Washington, Office of Education, Bureau of Research, 1970. (ERIC Document Reproduction Service No. ED 041 052)

Loevinger, J. A systematic approach to the construction and evalution of tests of ability. Psychological Monographs, 1947, 61, (4, Whole No. 285).

Macready, G. B. and Merwin, J. C. Homogeneity within item forms in domain-referenced testing. Educational and Psychological Measurement, 1973, 33, 351-360.

Mehrens, W. A. and Lehmann, Irvin J. Measurement and evaluation education and psychology. New York: Holt, Rinehart and Winston, 1978.

Millman, J. In W. J. Popham (Ed.), Evaluation in education: current applications. California: McCutchan Publishing Company, 1974.

Millman, J. Passing scores and test length for domain-referenced measurement. Review of Educational Research, 1973, 43, 205-216.

Nunnally, J. Introduction to psychological measurement. New York: McGraw-Hill Book Company, 1970.

Osburn, H. G. Item sampling for achievement testing. Educational and Psychological Measurement, 1968, 28, 95-104.

Pettie, A. A. and Oosterhof, A. C. Indices for item adequacy for individually administered mastery tests. Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, April 1976.

Popham, W. J. Educational Evaluation. Englewood Cliffs, N. J.: Prentice-Hall, 1975.

Popham, W. J. and Husek, T. R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.

Snedecor, G. W. and Cochran, W. G. Statistical methods. Ames, Iowa: Iowa State University Press, 1967.

Spady, W. G. Competency based education: Organizational Issues and Implications. Educational Researcher, 1977, 6, 9-14.

Stevens, S. Mathematics, measurement and psychophysics. In S. Stevens (Ed.), Handbook of Experimental Psychology. New York: Wiley, 1951.

Tyler, R. Basic principles of curriculum and instruction. Chicago: University of Chicago Press, 1970.

Appendix A

Review of Macready and Merwin (1973)

Ninety-one corpsmen from five randomly chosen youth conservation centers who were studying multiplication of whole numbers were subjects. Seventy-five items called for multiplication of whole numbers and were generated from 25 item forms. To produce the test, one item from each item form was generated and the 25 items were randomly ordered. This was done three times. The three sets of items were then administered in counterbalanced order. Time was allowed for all to finish each set. Each item was scored as right or wrong. Loevinger's Index of Homogeneity (Loevinger, 1947) was used as an index of the degree of homogeneity among items. The magnitude of these indexes varied considerably from one item form to another. Item forms with an index of homogeneity $\leq .50$ were considered relatively homogeneous, while those $> .50$ were considered relatively heterogeneous. In general, items from the same item form fell into the same cluster more often for relatively homogeneous item forms than was the case for relatively heterogeneous item forms.

In addition to homogeneity, the averages of the proportions of individuals getting each item within any item form correct, $\bar{P}$, was also examined. The $\bar{P}$ values for the relatively homogeneous item forms took on values

of intermediate magnitude with all except one of these values falling between .42 and .66. For relatively heterogeneous item forms, only two values fell between these limits.

Appendix B

Item Forms

1. Item Form One at the Knowledge Level in Diabetes

What is an example of an insulin with

- rapid / intermediate / prolonged action?
- an action peak of ½-1 / 1-2 / 2-8 / 8-12 / 12-14 hour(s) and
- duration of 5-7 / 12-16 / 24-28 / 36+ hour(s)?

2. Item Form Two at the Knowledge Level in Diabetes

Which of the folloiwng should you do when calculating

- the ideal body weight for a female / male with a small / medium / large frame?
- basal calories for an adult?
- activity calories for someone engaged in sedentary / moderate / strenuous activity?

3. What is the number of grams of carbohydrates, protein and

fat in one [fruit / vegetable / fat / whole milk / skim milk / bread / high-fat meat / medium-fat meat / low-fat meat] exchange according

to [Dr. Davidson's / The American Dietetic Association's]

exchange lists?

4. Item Form Four at the Knowledge Level in Diabetes

Which of the following is the [cause of / symptom of / appropriate terminology associated with]

[insulin shock? / diabetic coma? / having too much] [insulin / glucose] for the amount

of [glucose? / insulin?]

5. Item Form Five at the Knowledge Level in Diabetes

Which of the following is equal to one

[fruit
vegetable
fat
skim milk
bread
high-fat meat
medium-fat meat
low-fat meat
meat]

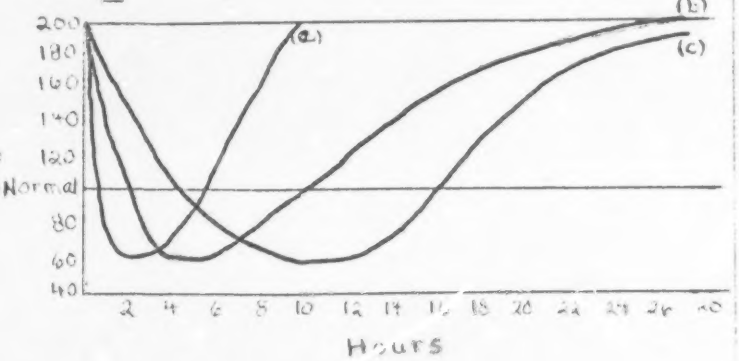
exchange using

[Dr. Davidson's
The American Dietetic Association's]

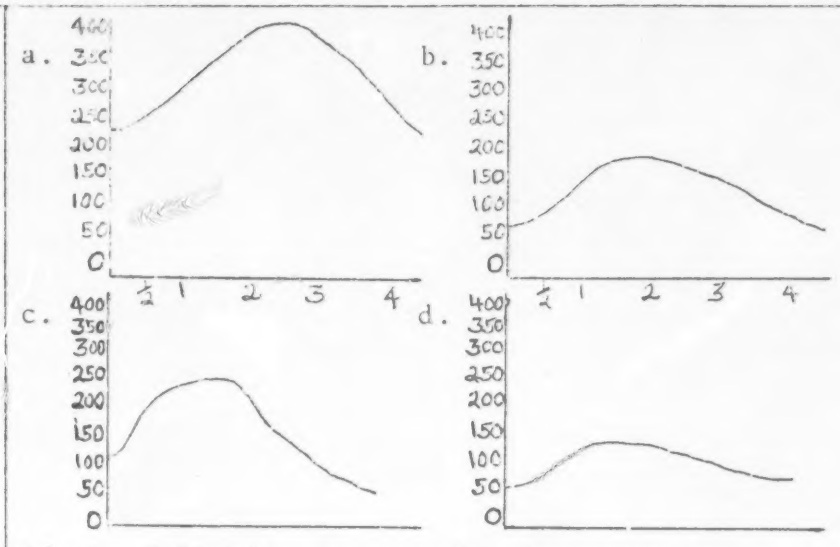
exchange lists?

6. Item Form One at the Comprehension Level in Diabetes

Which [type of insulin] is depicted by line [a / b / c] in the following graph?



[graph represents the results of a glucose tolerance test in which the person's glucose level has not returned to the fasting level within or up to three hours after administration of a glucose load and indicates the person as having diabetes?]



of the following represents the concept

of [U-40 / U-80 / U-100] in common terms?

7. Item Form Two at the Comprehension Level in Diabetes

Which of the following is [very important / not crucial] in the treatment of an [obese diabetic patient who does require insulin? / insulin-dependent non-obese diabetic patient?]

8. Item Form Three at the Comprehension Level in Diabetes

One of the early signs of diabetes may be

- polyuria defined as the excessive discharge of urine.
- fatigue.
- ketosis or the accumulation of ketone bodies.
- glycosuria or the presence of glucose in the urine.
- polydipsia defined as excessive thirst.
- blurring of vision.
- weight loss.

What is the primary cause of this initial manisfestation of diabetes?

9. Item Form One at the Application Level in Diabetes

What [is / are] the [ideal body weight / caloric requirements] for a [male / female] with a [small / medium / large] frame who is [5' / 5'5" / 5'8" / 6'], weighs [200 / 150] pounds and is engaged in [sedentary / moderate / strenuous] activity?

What [is / are] the number of grams of carbohydrate, protein and fat for a [1000 / 1500 / 2000] calorie diabetic diet with a distribution of 20% protein, [40% / 50%] carbohydrate and [40% / 30%] fat?

10. Item For Two at the Application Level in Diabetes

Which of the following menus would be appropriate for a [breakfast / lunch / dinner / snack]

consisting of [1-3] [low- / medium- / high-] fat meat exchanges, [0-1] [whole / skim] milk

exchanges, [0-1] fruit exchanges, [1-3] bread exchanges, [0-2]

vegetable exchanges, and [1-3] fat exchanges according to

[Dr. Davidson's / The American Dietetic Association's] exchange lists?

11. Item Form One at the Knowledge Level in Renal Disease

What is the normal range for
- serum: sodium? / potassium? / calcium? / chloride? / phosphorus? / creatinine?
- blood: urea nitrogen? / glucose?
- hematocrit / hemoglobin for a male? / female?

12. Item Form Two at the Knowledge Level in Renal Disease

Why would
- aluminum hydroxide preparations
- Kytakeral
- Kay exalate
- Inferon
- folic acid
- vitamin D preparations
- Aldemet
- heparin

be given to a patient with renal disease?

13. Item Form Three at the Knowledge Level in Renal Disease

Why would / wouldn't a patient with renal failure be able to use
- Hycal?
- Controlyte?
- Polycose?
- Wheat starch flour?
- Paygel Baking Mix?
- a salt substitute?
- Jello?
- Amin-Aid?

14. Item Form Four at the Knowledge Level in Renal Disease

How many grams of protein and milligrams of sodium and potassium does one [meat / bread / fat / wheat starch product / fruit / vegetable] in Group [A / B] exchange contain in the renal exchange lists?

15. Item Form Five at the Knowledge Level in Renal Disease

Which of the following characterizes [the nephrotic syndrome? / nephrosclerosis? / pyelonephritis? / polycystic kidney disease? / acute renal failure? / chronic glomerulonephritis?]

16. Item Form One at the Comprehension Level in Renal Disease

Renal patients who require dialysis often experience problems such as hypertension, leg cramps, anemia, neuropathy, pericarditis, arrhythmias, hypotension and itching.
What may be the [cause / symptoms] of

[leg cramps? / anemia? / hypotension? / itching? / neuropathy? / pericarditis? / arrhythmias? / hypertension?]

17. Item Form Two at the Comprehension Level in Renal Disease

The kidneys are responsible for

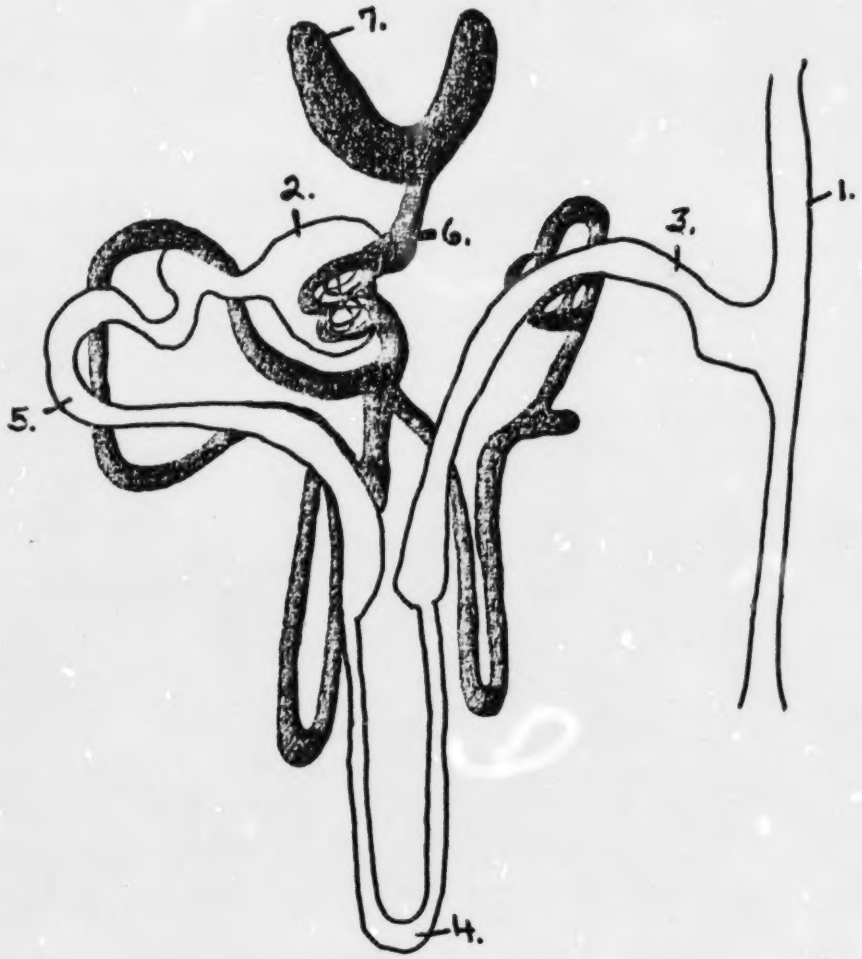
- the conversion of inactive Vitamin D.
- maintenance of the body's water balance.
- adjustment of the body's electrolyte balance.
- the production of erythropoietin.
- maintenance of acid-base balance.
- taking care of the amount of urea, uric acid and creatinine.

Which of the following describes this function of the kidney in broad terms?

18. Item Form Three at the Comprehension Level in Renal Disease

What is the part of the kidney labeled number [1 2 3 4 5 6 7]

in the following diagram?



19. Item Form One at the Application Level in Renal Disease

Which of the following would be the number and type of exchanges to approximately equal a [40-60 gram protein, [1-2 gram sodium [1.5-2 gram potassium diet?

20. Item Form Two at the Application Level in Renal Disease

If a renal patient was to have the following exchanges for a meal, 2 meat exchanges, 1 bread exchange, 4 fat exchanges, [1-2 [A- B- fruit exchanges and [1-2 [A- B- vegetable exchanges, which of the following would be the appropriate menu?

Appendix C

Tests

TEST ON DIABETES

1. Which of the following is equal to 1 fat exchange using Dr. Davidson's exchange lists?

   A. 1/8 of an avocado
   B. 1 tablespoon margarine
   C. 2 teaspoons mayonnaise
   D. 1 tablespoon light cream

2. Which of the following menus would be appropriate for a lunch meal consisting of 2 medium-fat meat exchanges, 2 bread exchanges, 2 vegetable exchanges and 3 fat exchanges according to The American Dietetic Association's exchange lists?

   A. 2 ounces boiled ham, ½ cup mashed potatoes, 1 small biscuit with 1 teaspoon butter, 1 cup broccoli with 1 teaspoon margarine
   B. ½ cup tuna with 3 teaspoons mayonnaise, 1 cup green beans, 6 saltines, 2 graham crackers
   C. 1 sandwich with 1 ounce bologna, 1 ounce cheddar cheese, 2 slices whole wheat bread, 1 lettuce leaf, 1 slice tomato and 3 teaspoons mayonnaise
   D. 2 ounces fresh trout, 1 small baked potato with 1 tablespoon sour cream, ½ cup baked beans, 1 cup asparagus with 1 teaspoon butter

3. Which of the following is the appropriate terminology when a diabetic patient has too much insulin for the amount of glucose?

   A. hypoinsulinemia
   B. hyperglycemia
   C. diabetic coma
   D. insulin shock

4. Which of the following is equal to 1 fruit exchange using Dr. Davidson's exchange lists?

   A. ½ cup blackberries
   B. 1 cup grapefruit juice
   C. 1 cup strawberries
   D. ¼ cup pineapple juice

5. Which of the following menus would be appropriate for a breakfast meal consisting of 1 meat exchange, 1 whole milk exchange, 1 fruit exchange, 1 fat exchange and 2 bread exchanges according to Dr. Davidson's exchange lists?

   A. 1 boiled egg, ½ cup grits with 1 teaspoon butter, ½ cup orange juice, 1 cup whole milk, 1 slice toast
   B. ½ grapefruit, 1 cup dry cereal with 1 cup whole milk, 1 ounce cheese on 1 slice toast with 1 teaspoon margarine
   C. 1 strip bacon, 1 cup whole milk, 2 slices toast with 1 teaspoon butter, ¼ cup grape juice
   D. 1 egg scrambled in 1 teaspoon margarine, ½ cup oatmeal, 1 slice toast with 1 teaspoon butter, 1/3 cup pineapple juice, 1 cup whole milk

6. Which of the following is equal to 1 low-fat meat exchange according to the American Dietetic Association's exchange lists?

   A. 1 ounce cheddar cheese
   B. ½ cup canned tuna
   C. 1 ounce shrimp
   D. 1 ounce veal breast

7. What is the number of grams of carbohydrate and protein in 1 skim milk exchange according to Dr. Davidson's exchange lists?

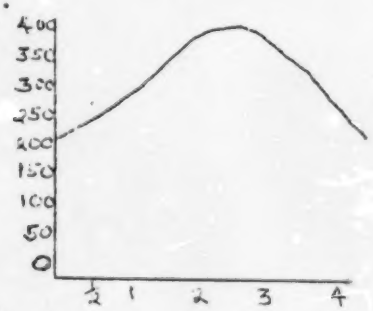
   A. 14 grams carbohydrate, 8 grams protein
   B. 14 grams carbohydrate, 6 grams protein
   C. 12 grams carbohydrate, 8 grams protein
   D. 12 grams carbohydrate, 6 grams protein

8. Which of the following should you do when calculating the ideal body weight for a female with a small frame?

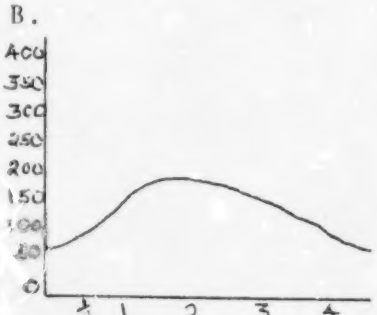
   A. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch
   B. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 5%
   C. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 10%
   D. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; add 10%

9. Which graph represents the results of a glucose tolerance test in which the person's glucose level has not returned to the fasting level within or up to three hours after administration of a glucose load and indicates the person as having diabetes?

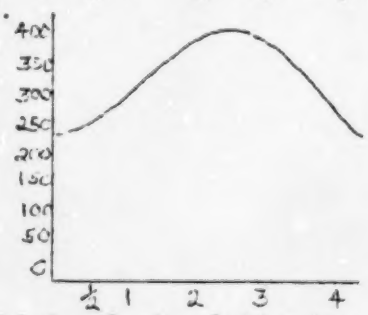
A.



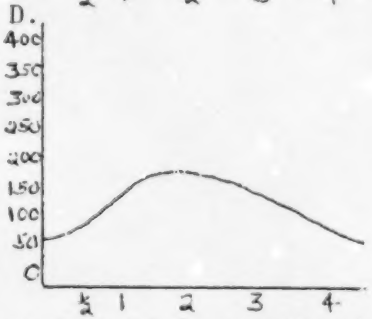
B.



C.



D.



10. Which of the following menus would be appropriate for a dinner meal consisting of 3 low-fat meat exchanges, 3 bread exchanges, 2 vegetable exchanges, 1 fruit exchange, 2 fat exchanges, and 1 skim milk exchange according to The American Dietetic Association's exchange lists?

A. 3 ounces rib roast, ½ cup rice with 1 teaspoon margarine, 2 small corn muffins with 1 teaspoon margarine, ½ cup spinach, ½ cup cauliflower, 1 small apple, 1 cup skim milk

B. 15 boiled shrimp, 1 small ear of corn with 1 teaspoon butter, a tossed salad with 1 cup lettuce, ½ cup tomatoes, ½ cup cucumbers and 1 tablespoon french dressing; 2 rolls with 1 teaspoon butter, 3/4 cup fresh strawberries, 1 cup skim milk

C. 3 ounces chicken breast with skin removed, ½ cup carrots with 1 teaspoon margarine, ½ cup brussel sprouts with 1 teaspoon margarine, 3 biscuits, 1 cup skim milk, ¼ of a medium honeydew
D. 3 ounces veal cutlet, ½ cup green peas with 1 teaspoon butter, 2/3 cup corn with 1 teaspoon butter, 1 cup tomato juice, ½ small banana, 1 cup skim milk

11. What is the number of grams of carbohydrate, protein and fat for a 2000 calorie diabetic diet with a distribution of 20% protein, 50% carbohydrate and 30% fat?

A. 100 grams protein, 250 grams carbohydrate, 67 grams fat
B. 75 grams protein, 187 grams carbohydrate, 50 grams fat
C. 50 grams protein, 125 grams carbohydrate, 33 grams fat
D. 100 grams protein, 100 grams carbohydrate, 22 grams fat

12. Which of the following should you do when calculating the ideal body weight for a female with a large frame?

A. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch
B. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 5%
C. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 10%
D. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; add 10%

13. Which of the following is equal to 1 fruit exchange according to The American Dietetic Association's exchange lists?

A. ½ cup blackberries
B. 1 cup grapefruit juice
C. 1 cup strawberries
D. ¼ cup pineapple juice

14. What is an example of an insulin with an action peak of 8 to 14 hours and duration of 36 plus hours?

A. NPH
B. Ultralente
C. Protamine Zinc
D. Semilente

15. Which of the following represents the concept of U-80 in common terms?

A. the type of insulin
B. the manufacturer of the insulin
C. the combination of insulin
D. the concentration of insulin

16. What is an example of an insulin with prolonged action?

A. Regular
B. Protamine Zinc
C. NPH
D. Lente

17. Which of the following should you do when calculating activity calories for someone engaged in sedentary activity?

A. allow 30%
B. allow 50%
C. allow 75%
D. allow 100%

18. What is an example of an insulin with an action peak of 1 to 2 hours and duration of 12-16 hours?

A. Semilente
B. Protamine Zinc
C. NPH
D. Regular

19. What is the number of grams of carbohydrate and protein in 1 vegetable exchange according to The American Dietetic Association's exchange lists?

A. 5 grams carbohydrate, 3 grams protein
B. 5 grams carbohydrate, 2 grams protein
C. 5 grams carbohydrate, 1 gram protein
D. 5 grams carbohydrate, .5 grams protein

20. Which of the following is very important in the treatment of an insulin dependent non-obese diabetic patient?

A. to decrease calories by restricting the amount of carbohydrate
B. to protect or improve beta cell function in order to improve glucose tolerance
C. to provide a daily consistency in the ratios of carbohydrate, protein and fat
D. to maintain the patient below ideal body weight for better control of blood glucose

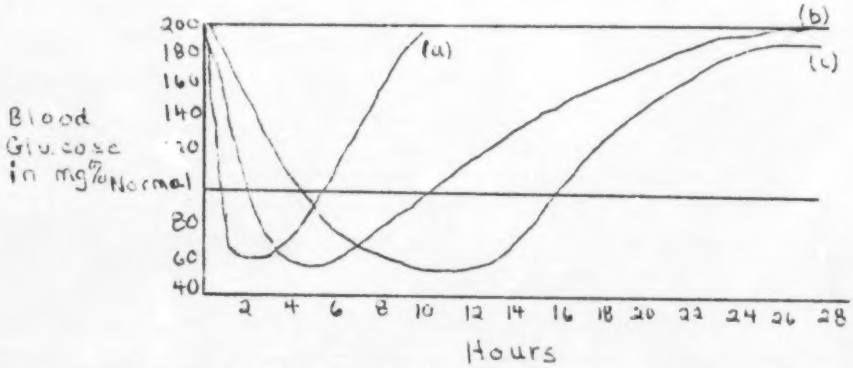
21. Which of the following is a symptom of having too much glucose for the amount of insulin?

A. pallor
B. sweating
C. blurred and double vision
D. extreme thirst

22. One of the early symptoms of diabetes may be polyuria defined as the excessive discharge of urine. What is the primary cause of this initial manifestation of diabetes?

A. a blood glucose level of approximately 160 mg% or greater
B. the excretion of large amounts of glucose in the urine resulting in loss of potential energy
C. the result of body water being utilized for excretion of glucose
D. the incomplete metabolism of fatty acids caused by inadequate carbohydrate utilization

23. What type of insulin is depicted by line c in the following graph?



A. Protamine Zinc with an action peak of 8-12 hours and duration of 24-28 hours
B. Regular with an action peak of ½-1 hour and duration of 5-7 hours
C. Ultralente with an action peak of 8-14 hours and duration of 364 hours
D. NPH with an action peak of 2-8 hours and duration of 24-28 hours

24. What is the number of grams of protein and fat in 1 low-fat meat exchange according to The American Dietetic Association's exchange lists?

A. 7 grams protein, 3 grams fat
B. 7 grams protein, 5 grams fat
C. 7 grams protein, 5.5 grams fat
D. 7 grams protein, 6 grams fat

25. What is an example of an insulin with rapid action?

A. Lente
B. Protamine Zinc
C. NPH
D. Regular

26. What is the number of grams of protein and fat in 1 high-fat meat exchange according to The American Dietetic Association's exchange lists?

A. 7 grams protein, 8 grams fat
B. 7 grams protein, 7 grams fat
C. 7 grams protein, 6 grams fat
D. 7 grams protein, 5 grams fat

27. Which of the following is a symptom of a diabetic patient having too much insulin for the amount of glucose?

A. extreme thirst
B. blurred and double vision
C. vomiting
D. flushed face

28. What is the number of grams of carbohydrate and protein in 1 skim milk exchange according to The American Dietetic Association's exchange lists?

A. 12 grams carbohydrate, 5 grams protein
B. 12 grams carbohydrate, 6 grams protein
C. 12 grams carbohydrate, 7 grams protein
D. 12 grams carbohydrate, 8 grams protein

29. Which of the following is a cause of insulin shock?

A. failing to take insulin
B. taking too much insulin
C. eating too much food
D. exercising too little

30. What are the caloric requirements for a male with a medium frame who is 5 feet, 6 inches tall, weighs 200 pounds and is engaged in moderate activity?

A. 1540
B. 1630
C. 1846
D. 2130

31. Which of the following is equal to 1 vegetable exchange using The American Dietetic Association's exchange lists?

A. 1 cup lettuce
B. ½ cup lima beans
C. ½ cup string beans
D. ½ cup winter squash

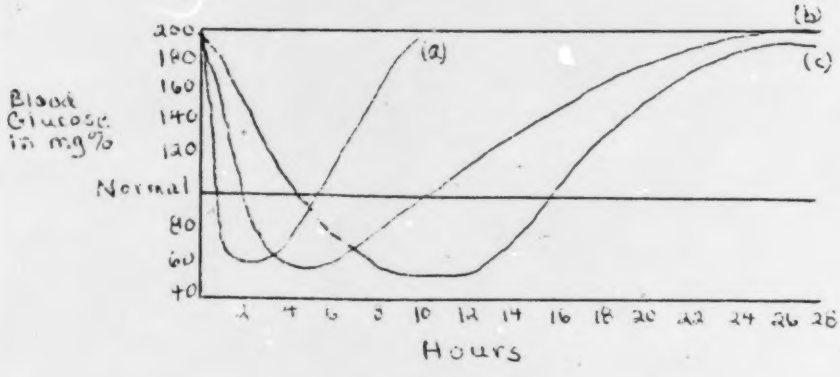
32. What is an example of an insulin with an action peak of 8-12 hours and duration of 36 plus hours?

A. Regular
B. Protamine Zinc
C. NPH
D. Lente

33. Which of the following should you do when calculating activity calories for someone engaged in moderate activity?

A. allow 30%
B. allow 50%
C. allow 75%
D. allow 100%

34. Which type of insulin is depicted by line a in the following graph?



A. Protamine Zinc with an action peak of 8-12 hours and duration of 24-28 hours
B. Regular with an action peak of ½-1 hour and duration of 5-7 hours
C. Ultralente with an action peak of 8-14 hours and duration of 36 plus hours
D. NPH with an action peak of 2-8 hours and duration of 24-28 hours

35. Which of the following is equal to 1 medium-fat exchange using the American Dietetic Association's exchange lists?

A. 1 ounce liver
B. 1 frankfurter
C. 1 ounce rib steak
D. 1 ounce bologna

36. Which of the following is the appropriate terminology when a diabetic patient has too much glucose for the amount of insulin?

A. insulin shock
B. diabetic coma
C. hypoglycemia
D. insulin reaction

37. Which of the following should you do when calculating the ideal body weight for a male with a small frame?

A. allow 106 pounds for the first 5 feet plus 6 pounds for each additional inch
B. allow 106 pounds for the first 5 feet plus 6 pounds for each additional inch; subtract 10%
C. allow 106 pounds for the first 5 feet plus 6 pounds for each additional inch; add 5%
D. allow 106 pounds for the first 5 feet plus 6 pounds for each additional inch; add 10%

38. Which of the following menus would be appropriate for a snack consisting of 1 meat exchange, 1 fat exchange, 1 bread exchange and 1 fruit exchange according to Dr. Davidson's exchange lists?

A. 1 ounce cheddar cheese, 5 saltines, ½ cup orange juice
B. ½ meat sandwich with 1 slice bread, 1 ounce turkey, 1 teaspoon mayonnaise, 1/3 cup apple juice

C. 3 graham crackers, 2 tablespoons peanut butter, 1 small apple
D. 1 carton plain yogurt with 1 cup fresh strawberries, 2 graham crackers with 1 tablespoon peanut butter

39. Which of the following is equal to 1 bread exchange using The American Dietetic Association's exchange lists?

A. ½ cup sweet potato
B. ½ cup green peas
C. 1 cup popcorn
D. 5 saltines

40. What is the ideal body weight for a female with a medium frame who is 5 feet, 5 inches tall?

A. 137 pounds
B. 136 pounds
C. 125 pounds
D. 112 pounds

41. What is an example of an insulin with an action peak of ½-1 hour and duration of 5-7 hours?

A. Lente
B. Protamine Zinc
C. NPH
D. Regular

42. Which of the following is very important in the treatment of an obese diabetic patient who does not require insulin?

A. to protect or improve beta cell function
B. to provide extra food for unusual exercise
C. to have consistency in the timing of meals
D. to provide consistency in the amount of fat

43. What is the number of grams of carbohydrate in 1 fruit exchange according to The American Dietetic Association's exchange lists?

A. 20 grams carbohydrate
B. 15 grams carbohydrate
C. 10 grams carbohydrate
D. 5 grams carbohydrate

44. Which of the following is a symptom of diabetic coma?

A. pallor
B. sweating
C. blurred and double vision
D. extreme thirst

45. What is the number of grams of protein and fat in 1 meat exchange according to Dr. Davidson's exchange lists?

A. 7 grams protein, 3 grams fat
B. 7 grams protein, 5 grams fat
C. 7 grams protein, 5.5 grams fat
D. 7 grams protein, 6 grams fat

46. Which of the following should you do when calculating the ideal body weight for a female with a medium frame?

A. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch
B. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 5%
C. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; subtract 10%
D. allow 100 pounds for the first 5 feet plus 5 pounds for each additional inch; add 10%

47. Which of the following is equal to 1 bread exchange using Dr. Davidson's exchange lists?

A. ½ cup sweet potato
B. ½ cup green peas
C. 1 cup popcorn
D. 5 saltines

48. What is an example of an insulin with intermediate action?

A. NPH
B. Ultralente
C. Protamine Zinc
D. Semilente

49. One of the early symptoms of diabetes may be glycosuria or the presence of glucose in the urine. What is the primary cause of this initial manifestation of diabetes?

A. a blood glucose level of approximately 160 mg% or greater
B. the excretion of large amounts of glucose in the urine
C. the result of body water being utilized for excretion of glucose
D. the incomplete metabolism of fatty acids caused by inadequate carbohydrate utilization

50. Which of the following is a symptom of insulin shock?

A. extreme thirst
B. blurred and double vision
C. vomiting
D. flushed face

51. Which of the following is not crucial in the treatment of an obese patient who does not require insulin?

A. to protect or improve beta cell function
B. to use food to treat, abort or prevent hypoglycemia
C. to provide a daily consistency in the number of calories
D. to decrease calories in order to attain ideal body weight

52. What is the number of grams of fat in 1 fat exchange according to The American Dietetic Association's exchange lists?

A. 10 grams fat
B. 7 grams fat
C. 5 grams fat
D. 3 grams fat

53. One of the early symptoms of diabetes may be blurring of vision. What is the primary cause of this initial manifestation of diabetes?

A. the result of body water being utilized for excretion of glucose
B. the increased glucose content of the lens
C. the excretion of large amounts of glucose in the urine resulting in loss of potential energy
D. the amount of glucose filtered by the glomeruli exceeds the reabsorptive capacity of the tubules

54. Which of the following is a cause of diabetic coma?

A. taking too much insulin
B. exercising too much
C. failing to take insulin
D. not eating enough food

55. What are the caloric requirements for a female with a large frame who is 5 feet, 5 inches tall, weighs 150 pounds and is engaged in sedentary activity?

A. 1875
B. 1787
C. 1560
D. 1287

56. Which of the following should you do when calculating activity calories for someone engaged in strenuous activity?

A. allow 30%
B. allow 50%
C. allow 75%
D. allow 100%

57. One of the early symptoms of diabetes may be polydipsia defined as excessive thirst. What is the primary cause of this initial manifestation of diabetes?

A. the result of body water being utilized for excretion of glucose
B. the increased glucose content of the lens
C. the excretion of large amounts of glucose in the urine resulting in loss of potential energy
D. the amount of glucose filtered by the glomeruli exceeds the reabsorptive capacity of the tubules

58. Which of the following is not crucial in the treatment of an insulin dependent non-obese diabetic patient?

A. to provide extra food for unusual exercise
B. to have consistency in the timing of meals
C. to protect or improve beta cell function
D. to use food to treat, abort or prevent hypoglycemia

59. Which of the following should you do when calculating basal calories for an adult?

A. allow 5 calories per pound of ideal body weight
B. allow 10 calories per pound of ideal body weight
C. allow 15 calories per pound of ideal body weight
D. allow 20 calories per pound of ideal body weight

60. What is an example of an insulin with an action peak of 2-8 hours and duration of 24-28 hours?

A. NPH
B. Ultralente
C. Protamine Zinc
D. Semilente

TEST ON RENAL DISEASE

1. Which of the following would be the number and type of exchanges to approximately equal a 50 gram protein, 1.5 gram sodium, 2 gram potassium diet?

   A. 7 meat exchanges, 3 B-vegetable exchanges, 3 bread exchanges, 2 B-fruit exchanges, 15 fat exchanges, 1/8 teaspoon salt
   B. 7 meat exchanges, 3 bread exchanges, 1 A-vegetable exchange, 2 B-vegetable exchanges, 3 B-fruit exchanges, 5 fat exchanges
   C. 4 meat exchanges, 2 B-vegetable exchanges, 3 bread exchanges, 3 B-fruit exchanges, 8 fat exchanges
   D. 6 meat exchanges, 2 bread exchanges, 3 B-vegetable exchanges, 2 B-fruit exchanges, 8 fat exchanges, 1/4 teaspoon salt

2. What is the normal range for serum calcium?

   A. 3.5 - 5.5 mEq/L
   B. 135 - 145 mEq/L
   C. 8 - 20 mg%
   D. 8.5 - 10.8 mg%

3. Why wouldn't a patient with renal failure be able to use a salt substitute?

   A. it contains potassium instead of sodium
   B. it contains half the amount of sodium as salt
   C. it contains a significant amount of chloride
   D. it contains one-fourth the amount of sodium as salt

4. Why would heparin be given to a patient with renal disease?

   A. to control hypertension
   B. to increase urine output
   C. to increase red cell production
   D. to prevent blood from clotting

5. The kidneys are responsible for the production of erythropoietin. Which of the following describe this function of the kidney in broad terms?

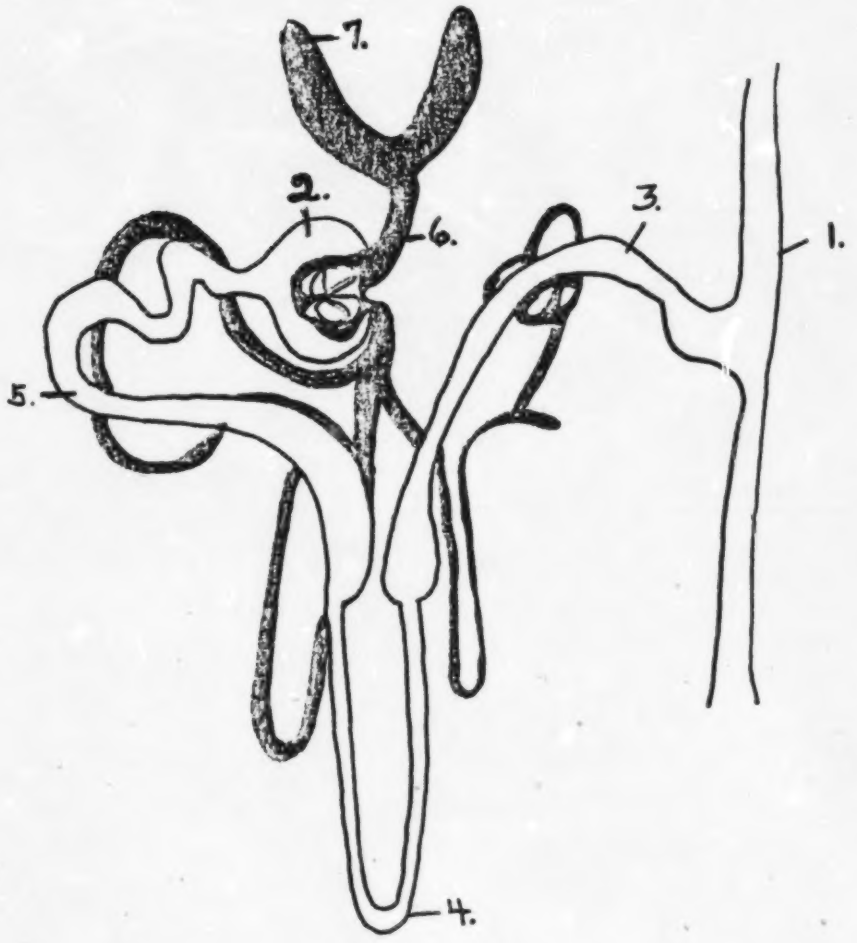
   A. excretory
   B. regulatory
   C. stimulatory
   D. endocrine

6. If a renal patient was to have the following exchanges for a meal, 2 meat exchanges, 1 bread exchange, 1 A-vegetable exchange, 1 B-vegetable exchange, 1 A-fruit exchange, 1 B-fruit exchange and 4 fat exchanges, which of the following would be the appropriate menu?

   A. 2 ounces steak, ½ cup rice with 1 teaspoon butter, ½ cup carrots with 1 teaspoon butter, ½ cup summer squash with 1 teaspoon butter, ½ cup apple juice and ½ cup peaches
   B. ½ cup water-packed tuna with 1 tablespoon mayonnaise, 3 unsalted soda crackers, ½ cup broccoli with 1 teaspoon butter, and 1 cup fresh strawberries
   C. 2 ounces chicken breast, 1 biscuit with 1 teaspoon butter, ¼ cup green peas with 1 teaspoon butter, ½ cup corn with 1 teaspoon butter, 1 cup frozen blueberries with 1 tablespoon heavy cream
   D. 2 ounces broiled hamburger patty, ¼ cup brussel sprouts with 1 teaspoon butter, ½ cup cabbage with 2 teaspoons mayonnaise, 1 cup grape juice, and 1 roll with 1 teaspoon butter

7. What is the part of the kidney labeled number 2 in the following diagram?

   A. Loop of Henle
   B. collecting tubule
   C. distal tubule
   D. Bowman's capsule



8. Which of the following characterizes polycystic kidney disease?

   A. gradual fibrosis of glomeruli and degeneration of the tubules
   B. recurrent infections of the kidney leading to scarring and contraction
   C. a defect present at birth which leads to renal failure later in life
   D. crystalline deposits of uric acid salts which leads to scarring and atrophy

9. Which of the following would be the number and type of exchanges to approximately equal a 60 gram protein, 1 gram sodium, 2 grams potassium diet?

   A. 7 meat exchanges, 3 B-vegetable exchanges, 3 bread exchanges, 2 B-fruit exchanges, 15 fat exchanges, 1/8 teaspoon salt
   B. 7 meat exchanges, 3 bread exchanges, 1 A-vegetable exchange, 2 B-vegetable exchanges, 3 B-fruit exchanges, 5 fat exchanges
   C. 4 meat exchanges, 2 B-vegetable exchanges, 3 bread exchanges, 3 B-fruit exchanges, 8 fat exchanges
   D. 6 meat exchanges, 2 bread exchanges, 3 B-vegetable exchanges, 2 B-fruit exchanges, 8 fat exchanges, ¼ teaspoon salt

10. Which of the following characterizes chronic glomerulonephritis?

   A. gradual fibrosis of glomeruli and degeneration of the tubules
   B. recurrent infections of the kidney leading to scarring and contraction
   C. a defect present at birth which leads to renal failure later in life
   D. crystalline deposits of uric acid salts which leads to scarring and atrophy

11. Why would hytakeral be given to a patient with renal disease?

   A. to increase calcium absorption
   B. to decrease potassium absorption
   C. to decrease blood urea nitrogen
   D. to decrease phosphorus absorption

12. Which of the following characterizes chronic renal failure?

   A. diminished renal reserve, renal insufficiency, and urea
   B. rapid, steadily increasing azotemia with or without oliguria
   C. renal arteriolar necrosis associated with hypertension
   D. hematuria, renal functional impairment and nerve deafness

13. What is the normal range for serum chloride?

A. .8 - 1.5 mg%
B. 60 - 110 mg%
C. 2.5 - 4.5 mg%
D. 95 - 110 mEq/L

14. Why would aldomet be given to a patient with renal disease?

A. to control hypertension
B. to increase urine output
C. to increase red cell production
D. to prevent blood from clotting

15. If a renal patient was to have the following exchanges for a meal, 2 meat exchanges, 1 bread exchange, 2 A-vegetable exchanges, 2 B-fruit exchanges and 4 fat exchanges, which of the following would be the appropriate menu?

A. 2 ounces steak, ½ cup rice with 1 teaspoon butter, ¼ cup carrots with 1 teaspoon butter, ½ cup summer squash with 1 teaspoon butter, ⅓ cup apple juice, and ½ cup peaches
B. ½ cup water-packed tuna with 1 tablespoon mayonnaise, 3 unsalted soda crackers, ½ cup broccoli with 1 teaspoon butter, and 1 cup fresh strawberries
C. 2 ounces chicken breast, 1 biscuit with 1 teaspoon butter, ¼ cup green peas with 1 teaspoon butter, ½ cup corn with 1 teaspoon butter, 1 cup frozen blueberries with 1 tablespoon heavy cream
D. 2 ounces broiled hamburger patty, ¼ cup brussel sprouts wtih 1 teaspoon butter, ½ cup cabbage with 2 teaspoons mayonnaise, 1 cup grape juice, and 1 roll with 1 teaspoon butter

16. What is the normal range for serum potassium?

A. 3.5 - 5.5 mEq/L
B. 135 - 145 mEq/L
C. 8 - 20 mg%
D. 8.5 - 10.8 mg%

17. The kidneys are responsible for maintenance of acid-base balance. Which of the following describes this function of the kidney in broad terms?

A. excretory
B. regulatory
C. stimulatory
D. endocrine

18. If a renal patient was to have the following exchanges for a meal, 2 meat exchanges, 1 bread exchange, 2 B-vegetable exchanges, 2 A-fruit exchanges and 4 fat exchanges, which of the following would be the appropriate menu?

A. 2 ounces steak, ½ cup rice with 1 teaspoon butter, ¼ cup carrots with 1 teaspoon butter, ½ cup summer squash with 1 teaspoon butter, ½ cup apple juice and ½ cup peaches
B. ½ cup water-packed tuna with 1 tablespoon mayonnaise, 3 unsalted soda crackers, ½ cup broccoli with 1 teaspoon butter, and 1 cup fresh strawberries
C. 2 ounces chicken breast, 1 biscuit with 1 teaspoon butter, ⅓ cup green peas with 1 teaspoon butter, ¼ cup corn with 1 teaspoon butter, 1 cup frozen blueberries with 1 tablespoon heavy cream
D. 2 ounces broiled hamburger patty, ¼ cup brussel sprouts with 1 teaspoon butter, ½ cup cabbage with 2 teaspoons mayonnaise, 1 cup grape juice, and 1 roll with 1 teaspoon butter

19. Renal patients who require dialysis often experience problems such as hypertension, leg cramps, anemia, neuropathy, pericarditis, arrhythmias, hypotension and itching. What may be the symptoms of pericarditis?

A. a persistent pain in the chest
B. results in non-noticable symptoms
C. numbness of feet and lower legs
D. edema, coughing or shortness of breath

20. Why would a patient with renal failure be able to use Polycose?

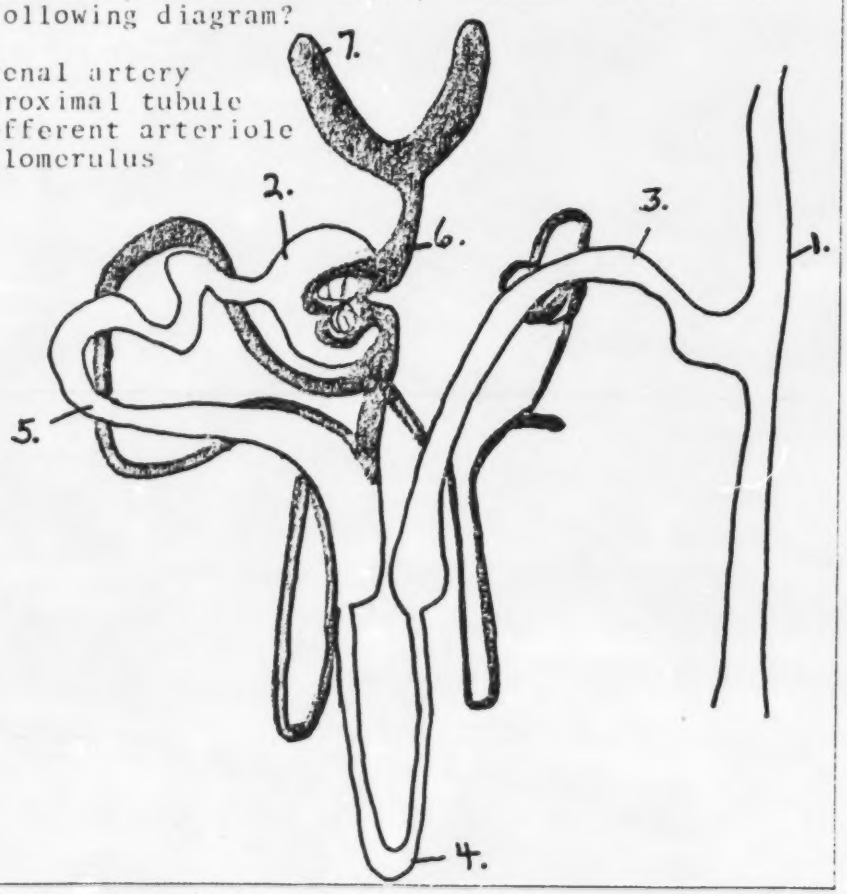
A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

21. How many grams of protein and milligrams of sodium and potassium does 1 vegetable exchange in Group B contain in the renal exchange lists?

A. 1.5 grams protein, 16 mg sodium, 219 mg potassium
B. 2 grams protein, 148 mg sodium, 38 mg potassium
C. 7 grams protein, 33 mg sodium, 103 mg potassium
D. .5 grams protein, 2 mg sodium, 96 mg potassium

22. What is the part of the kidney labeled number 6 in the following diagram?

A. renal artery
B. proximal tubule
C. efferent arteriole
D. glomerulus



23. How many grams of protein and milligrams of sodium and potassium does 1 fruit exchange in Group A contain in the renal exchange lists?

A. 1.5 grams protein, 16 mg sodium, 219 mg potassium
B. 2 grams protein, 148 mg sodium, 38 mg potassium
C. 7 grams protein, 33 mg sodium, 103 mg potassium
D. .5 grams protein, 2 mg sodium, 96 mg potassium

24. Why would a patient with renal failure be able to use Paygel Baking Mix?

A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

25. What is the normal range for serum sodium?

A. 3.5 - 5.5 mEq/L
B. 135 - 145 mEq/L
C. 8 - 20 mg%
D. 8.5 - 10.8 mg%

26. Which of the following characterizes acute glomerulonephritis?

A. diffuse inflammatory changes in glomeruli with albuminuria, hematuria, edema, hypertension and nitrogen retention
B. severely decreased urine output with a low specific gravity
C. albuminuria, hypoalbuminenia, hypercholesteremia, and often massive edema
D. sclerosis of the afferent arterioles followed by degenerative destruction of the nephrons

27. Why would a patient with renal failure be able to use wheat starch flour?

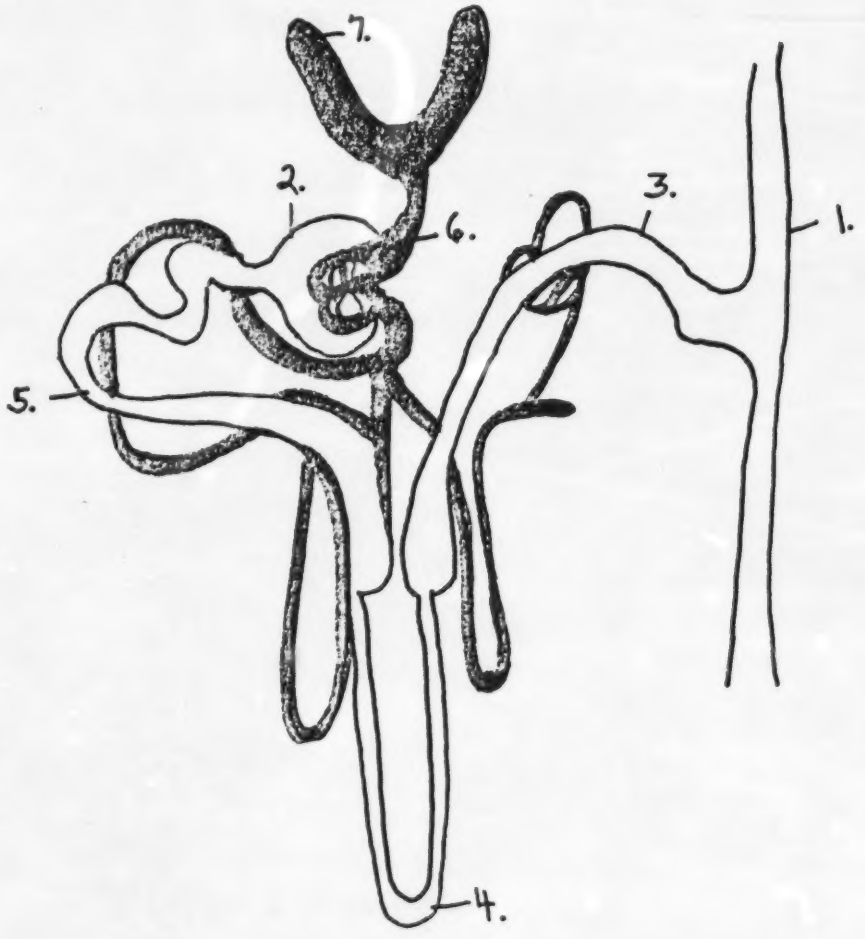
A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

28. What is the normal range for hematocrit for a female?

A. 14 - 18 mg%
B. 36 - 47 %
C. 12 - 16 mg%
D. 47 - 54 %

29. What is the part of the kidney labeled number 4 in the following diagram?

A. Loop of Henle
B. collecting tubule
C. distal tubule
D. Bowman's capsule



30. Why would kayexalate be given to a patient with renal failure?

A. to increase calcium absorption
B. to decrease potassium absorption
C. to decrease blood urea nitrogen
D. to decrease phosphorus absorption

31. Why would vitamin D preparations be given to a patient with renal failure?

A. to correct severe anemia
B. to maintain serum calcium levels
C. to replace loses which occur during dialysis
D. to decrease serum sodium levels

32. The kidneys are responsible for the conversion of biologically inactive vitamin D. Which of the following describe this function of the kidney in broad terms?

A. excretory
B. regulatory
C. stimulatory
D. endocrine

33. What is the noraml range for serum phosphorus?

A. .8 - 1.5 mg%
B. 60 - 110 mg%
C. 2.5 - 4.5 mg%
D. 95 - 110 mEq/L

34. Why would a patient with renal failure be able to use Hycal?

A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

35. How many grams of protein and milligrams of sodium and potassium does 1 vegetable exchange in Group A contain in the renal exchange lists?

A. 1.5 grams protein, 16 mg sodium, 127 mg potassium
B. a trace of protein, 49 mg sodium, 6 mg potassium
C. .5 grams protein, 2 mg sodium, 177 mg potassium
D. .75 grams protein, 148 mg sodium, 77 mg potassium

36. How many grams of protein and milligrams of sodium and potassium does 1 meat exchange contain in the renal exchange lists?

A. 1.5 grams protein, 16 mg sodium, 219 mg potassium
B. 2 grams protein, 148 mg sodium, 38 mg potassium
C. 7 grams protein, 33 mg sodium, 103 mg potassium
D. .5 grams protein, 2 mg sodium, 96 mg potassium

37. What is the normal range for hemoglobin for a male?

A. 14 - 18 mg%
B. 36 - 47 %
C. 12 - 16 mg%
D. 47 - 54 %

38. How many grams of protein and milligrams of sodium and potassium does 1 fat exchange contain in the renal exchange lists?

A. 1.5 grams protein, 16 mg sodium, 127 mg potassium
B. trace of protein, 49 mg sodium, 6 mg potassium
C. .5 grams protein, 2 mg sodium, 177 mg potassium
D. .75 grams protein, 148 mg sodium, 77 mg potassium

39. Which of the following characterizes nephrosclerosis?

A. diffuse inflammatory changes in glomeruli with albuminuria, hematuria, edema, hypertension and nitrogen retention
B. severely decreased urine output with a low specific gravity
C. albuminuria, hypoalbuminemia, hypercholesteremia, and often massive edema
D. sclerosis of the afferent arterioles followed by degenerative destruction of the nephrons

40. Renal patients who require dialysis often experience problems such as hypertension, leg cramps, anemia, neuropathy, pericarditis, arrhythmias, hypotension and itching. What may be the symptoms of hypertension?

A. a persistent pain in the chest
B. results in non-noticable symptoms
C. numbness of feet and lower legs
D. edema, coughing and shortness of breath

41. Which of the following would be the number and type of exchanges to approximately equal a 40 gram protein, 1 gram sodium, 1.5 gram potassium diet?

A. 7 meat exchanges, 3 B-vegetable exchanges, 3 bread exchanges, 2 B-fruit exchanges, 15 fat exchanges, 1/8 teaspoon salt
B. 7 meat exchanges, 3 bread exchanges, 1 A-vegetable exchange, 2 B-vegetable exchanges, 3 B-fruit exchanges, 5 fat exchanges
C. 4 meat exchanges, 2 B-vegetable exchanges, 3 bread exchanges, 3 B-fruit exchanges, 8 fat exchanges
D. 6 meat exchanges, 2 bread exchanges, 3 B-vegetable exchanges, 2 B-fruit exchanges, 8 fat exchanges, ¼ teaspoon salt

42. Which of the following characterizes acute renal failure?

A. diffuse inflammatory changes in glomeruli with albuminuria, hematuria, edema, hypertension and nitrogen retention
B. severely decreased urine output with a low specific gravity
C. albuminuria, hypoalbuminenia, hypercholesteremia, and often massive edema
D. sclerosis of the afferent arterioles followed by degenerative destruction of the nephrons

43. How many grams of protein and milligrams of sodium and potassium does 1 bread exchange contain in the renal exchange list?

A. 1.5 grams protein, 16 mg sodium, 219 mg potassium
B. 2 grams protein, 148 mg sodium, 38 mg potassium
C. 7 grams protein, 33 mg sodium, 103 mg potassium
D. .5 grams protein, 2 mg sodium, 96 mg potassium

44. Renal patients who require dialysis often expereince problems such as hypertension, leg cramps, anemia, neuropathy, pericarditis, arrhythmias, hypotension and itching. What may be the cause of leg cramps?

A. rapid shifts of salt and water
B. salt and water depletion
C. decreased production of erythropoietin
D. a chemical imbalance

45. Which of the following would be the number and type of exchanges to approximately equal a 60 gram protein, 2 gram sodium, 2 gram potassium diet?

A. 7 meat exchanges, 3 B-vegetable exchanges, 3 bread exchanges, 2 B-fruit exchanges, 15 fat exchanges, 1/8 teaspoon salt
B. 7 meat exchanges, 3 bread exchanges, 1 A-vegetable exchange, 2 B-vegetable exchanges, 3 B-fruit exchanges, 5 fat exchanges
C. 4 meat exchanges, 2 B-vegetable exchanges, 3 bread exchanges, 3 B-fruit exchanges, 8 fat exchanges
D. 6 meat exchanges, 2 bread exchanges, 3 B-vegetable exchanges, 2 B-fruit exchanges, 8 fat exchanges, ¼ teaspoon salt

46. Why would a patient with renal failure be able to use Controlyte?

A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

47. The kidneys are responsible for taking care of the amount of urea, uric acid and creatinine. Which of the following describes this function of the kidney in broad terms?

A. excretory
B. regulatory
C. stimulatory
D. endocrine

48. Why would folic acid be given to a patient with renal disease?

A. to correct severe bone disease
B. to maintain serum calcium levels
C. to replace loses which occur during dialysis
D. to decrease serum sodium levels

49. Renal patients who require dialysis often experience problems such as hypertension, leg cramps, anemia, neuropathy, pericarditis, arrhythmias, hypotension and itching. What may be the cause of hypotension?

A. rapid shifts of salt and water
B. salt and water depletion
C. decreased production of erythropoietin
D. a chemical imbalance

50. Which of the following characterizes the nephrotic syndrome?

A. diffuse inflammatory changes in glomeruli with albuminuria, hematuria, edema, hypertension and nitrogen retention
B. severely decreased urine output with a low specific gravity
C. albuminuria, hypoalbuminemia, hypercholesteremia, and often massive edema
D. sclerosis of the afferent arterioles followed by degenerative destruction of the nephrons

51. How many grams of protein and milligrams of sodium and potassium does 1 fruit exchange in Group B contain in the renal exchange list?

A. 1.5 grams protein, 16 mg sodium, 127 mg potassium
B. trace of protein, 49 mg sodium, 6 mg potassium
C. .5 grams protein, 2 mg sodium, 177 mg potassium
D. .75 grams protein, 148 mg sodium, 77 mg potassium

52. If a renal patient was to have the following exchanges for a meal, 2 meat exchanges, 1 bread exchange, 2 A-vegetable exchanges, 2 A-fruit exchanges, and 4 fat exchanges, which of the following would be the appropriate menu?

A. 2 ounces steak, ½ cup rice with 1 teaspoon butter, ½ cup carrots with 1 teaspoon butter, ½ cup summer squash with 1 teaspoon butter, ½ cup apple juice, and ½ cup peaches
B. ½ cup water-packed tuna with 1 tablespoon mayonnaise, 3 unsalted soda crackers, ½ cup broccoli with 1 teaspoon butter, and 1 cup fresh strawberries
C. 2 ounces chicken breast, 1 biscuit with 1 teaspoon butter, ¼ cup green peas with 1 teaspoon butter, ¼ cup corn with 1 teaspoon butter, 1 cup frozen blueberries with 1 tablespoon heavy cream
D. 2 ounces broiled hamburger patty, ¼ cup brussel sprouts with 1 teaspoon butter, ½ cup cabbage with 2 teaspoons mayonnaise, 1 cup grape juice, and 1 roll with 1 teaspoon butter

53. Why would inferon be given to a patient with renal disease?

A. to correct severe anemia
B. to maintain serum calcium levels
C. to replace loses occuring during dialysis
D. to decrease serum sodium levels

54. Why would a patient with renal failure be able to use Amin-Aid?

A. the majority of low-quality protein has been removed
B. it contains negligible amounts of electrolytes
C. it contains the essential amino acids
D. it provides significant amounts of calcium

55. Which of the following characterizes pyelonephritis?

A. gradual fibrosis of glomeruli and degeneration of the tubules
B. recurrent infections of the kidney leading to scarring and contraction
C. a defect present at birth which leads to renal failure later in life
D. crystalline deposits of uric acid salts which leads to scarring and atrophy

56. Why wouldn't a patient with renal failure be able to use jello?

A. it contains a significant amount of potassium
B. it contains protein of low biological value
C. it contains too much protein
D. it contains a large amount of sodium

57. How many grams of protein and milligrams of sodium and potassium does 1 wheat starch product exchange contain in the renal exchange list?

A. 1.5 grams protein, 16 mg sodium, 127 mg potassium
B. trace of protein, 49 mg sodium, 6 mg potassium
C. .5 grams protein, 2 mg sodium, 177 mg potassium
D. .75 grams protein, 147 mg sodium, 77 mg potassium

58. What is the normal range for blood urea nitrogen?

A. 3.5 - 5.5 mEq/L
B. 135 - 145 mEq/L
C. 8 - 20 mg%
D. 8.5 - 10.8 mg%

59. Why would aluminum hydroxide preparations be given to a patient with renal disease?

A. to increase calcium absorption
B. to decrease potassium absorption
C. to decrease blood urea nitrogen
D. to decrease phosphorus absorption

60. What is the part of the kidney labeled number 5 in the following diagram?

A. renal artery
B. proximal tubule
C. efferent arteriole
D. glomerulus